INTERPRETATIONAL AND INSTRUCTIONAL FACTORS IN THE SELECTION TASK

By

RICHARD D. PLATT

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1992

ACKNOWLEDGMENTS

I would like to acknowledge my committee chairperson, Professor Richard A. Griggs, for his help and guidance. I would also like to thank the other members of my committee: Professor C. Michael Levy, Professor Ira S. Fischler, Dr. M. Jeffrey Farrar, and Professor James J. Algina.

TABLE OF CONTENTS

LIST OF TABLES

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

INTERPRETATIONAL AND INSTRUCTIONAL FACTORS IN THE SELECTION TASK

By

Richard D. Platt

December 1992

Chairman: Richard A. Griggs
Major Department: Psychology

The present study examined performance on Wason's four-card abstract selection task, the most-researched task in reasoning. Baseline performance on this logical reasoning task, regardless of education level, is usually less than 10% correct. It is hypothesized that the two primary sources of difficulty are rule ambiguity and the lack of analytic processing. Four factorial experiments with 356 subjects were conducted to test this hypothesis. In Experiment 1, performance was improved by explicating the implication rule. The majority of subjects, however, still failed to make the correct selection. To increase analytic processing, subjects were required in Experiment 2 to provide reasons for their selection or nonselection of each of the cards. This response procedure paired with an explicated rule led to further improvements in performance (over 50% correct selections). Because the explicated rule

in the first two experiments prevented matching bias (a prevalent nonanalytic process) from operating, explicated rules that allowed and prevented matching bias were examined in Experiment 3. No effects for rule type were found, and facilitation on the task was still observed. In Experiment 4, the influence of the type of selection instruction (true-false vs. violation) was studied. Paired with an explicated rule and the reason-giving response format, violation instructions led to one of the highest levels of performance ever observed for the abstract selection task, 78% correct. Because of the importance of this result, it was replicated twice. The results for the four experiments are discussed with respect to the major theoretical approaches in the reasoning area, and mental models theory is found to offer the best explanation.

CHAPTER 1
INTRODUCTION

The Wason selection task has been one of the major deductive reasoning research paradigms for the past two decades. Dozens of papers have been written examining different versions of the same general problem type since its first introduction by Wason in 1966. In its original abstract form, the selection task presents the subject with four cards and a conditional rule that refers to these cards. The cards have a letter on one side and a number on the other side. The subject can only see one side of each card and is asked to indicate those cards that must be turned over in order to see whether the conditional rule is true or false. For example, one rule that has been used is "If a card has a vowel on one side, then it has an even number on the other side." The four cards would then show a vowel, a consonant, an even number, and an odd number. If the "If..., then..." rule is interpreted as a material implication, then we can represent the form of the argument as P-->Q. The cards then represent the four logical cases: P, NOT-P, Q, and NOT-Q.

In deciding which cards must be investigated further to determine the truth value of the rule, most subjects select either P alone or P & Q (Johnson-Laird & Wason, 1970a).

However, the logically correct response would be to check P and NOT-Q. P must be checked because the presence of NOT-Q on the other side would falsify the rule. NOT-Q must be checked because there might be P on the other side which would likewise falsify the rule. Q need not be checked because the rule is true for this card regardless of whether P or NOT-P is on the other side. Therefore, subjects tend to make two types of mistakes. They erroneously include Q in their selections and fail to include NOT-Q. A few subjects make the first of these errors but not the second. They select P, Q, & NOT-Q and this selection has been referred to as partial insight (Johnson-Laird & Wason, 1970a).

Much of the literature on the selection task has focused on the finding that certain concrete thematic contents can facilitate performance (for a review, see Griggs, 1983). The focus of this line of research has been to provide an explanation of why facilitation occurs in these cases rather than why facilitation does not occur in more abstract problems. Among the mechanisms proposed to explain this facilitation have been the memory-cueing hypothesis (Griggs & Cox, 1982), pragmatic reasoning schemas (Cheng & Holyoak, 1985), and Darwinian algorithms (Cosmides, 1989). Literally, dozens of different problems have been designed in the quest for support of these different theoretical accounts. Some of these problems have changed so much that they hold little outward resemblance to the

original selection task (for example, see Cosmides's unfamiliar social contract problems, 1989).

The present study will re-examine the original abstract task in an attempt to determine the sources of difficulty for this problem. It will be argued that central to the difficulty of the selection task is ambiguity in interpretation of the rule and lack of analytic processing. In the first section of this introduction, the basic research regarding the interpretation of If P, Q statements will be reviewed. The second section introduces the major theoretical explanations of observed selection task performance focusing primarily on explanations of the standard abstract problem. In order to determine why the problem is difficult, it is important to understand what processes the subjects may be using to arrive at the answers that they give, even if those answers are incorrect. For example, if short-cut strategies are readily available, then subjects may opt for the easy solution rather than trying to use more sophisticated analytic tools. The third section reviews the fairly extensive research on attempts to facilitate performance for the abstract selection task. These attempts have, for the most part, resulted in marginal improvements in performance. The final topic to be considered is the presence of powerful instructional and format variables. A number of studies using abstract and thematic materials have shown an influence of these types of

variables in determining not only the level of facilitation, but also the type of error pattern that is produced.

Possible Interpretations of the Rule

In most of the research on the selection task, the rule has been presented in an If P, Q form. Therefore, discussion about logical performance in the selection task has been predicated on the assumption that subjects interpret the If P, Q rule as a material implication. The logical relationship of material implication is considered true in all cases except where the antecedent is true and the consequent is false. There are two valid forms of argument derived from the material implication P-->Q: Modus ponens holds that given P is true, one can conclude that Q is true, and modus tollens holds that given Q is false, one can conclude that P is false. There are also two common fallacies: Denying the antecedent states that given P is false, one can conclude that Q is also false, while affirming the consequent states that given Q is true, one can conclude that P is true. Both of these last two arguments are logically unwarranted for material implication.

Fillenbaum (1975) criticizes psychological research for assuming that If P, Q statements have a single meaning (material implication) that is not influenced by the content of P and Q. He found that subjects gave different types of paraphrases when presented with If P, Q sentences expressing a variety of relationships such as causation, conditional

promises, conditional threats, temporal conditionals, contingent universals, and class inclusions. Additionally, subjects described as strange or extraordinary those sentences in which the antecedent and consequent were not topically related. This describes the situation in the abstract selection task where the rule arbitrarily links abstract symbols such as letters and numbers.

Geis and Zwicky (1971) argue that there are "invited inferences" present in many conditional statements. For example, they argue that "If you mow the lawn, I'll give you five dollars" invites the logical fallacy of denying the antecedent. They claim that this type of inference applies to promises, threats, law-like statements, commands, and counterfactual conditionals. Rumain, Connell, and Braine (1983) tested the role of invited inferences in determining the conclusions subjects would draw in a conditional syllogism task. They found that when the major premise explicitly countermanded the invited inferences, subjects were less likely to endorse the fallacies of denying the antecedent and affirming the consequent. However, they were also less likely to endorse the logically correct modus tollens. This finding provides evidence in support of invited inferences, although explicitly uninviting the inferences still does not produce increased logically correct performance.

Wason (1966) has claimed that subjects adopt a three-value logic in evaluating If P, Q statements. According to

the logical rules governing the material implication P-->Q: P, Q is true; P, NOT-Q is false; NOT-P, Q is true; and NOT-P, NOT-Q is true. However, Wason has claimed that subjects judge these latter two cases as irrelevant. In so doing, they introduce the third value of irrelevancy into their logic. It should be noted that this interpretation of the If P, Q statement would lead to the same selections as the material implication interpretation for the selection task. Wason argued that these selections do not occur because the subject is only looking for instances that verify the rule.

Johnson-Laird and Tagart (1969) considered four different ways of expressing material implication in natural language and asked subjects to produce truth tables for each. Subjects were allowed to use three truth values: true, false, and irrelevant. They found that subjects' response pattern for If P, Q statements followed Wason's predictions. The majority (19/24) of subjects judged both cases containing NOT-P as irrelevant.

Taplin (1971) had subjects perform a task similar to Johnson-Laird and Tagart (1969) but they were limited in their responses to a two-value logic (true or false). He found that only 45 percent of the subjects responded in a consistent truth functional pattern. The vast majority of those who responded consistently gave a biconditional interpretation. Taplin and Staudenmayer (1973) replicated the predominance of biconditional interpretations among

those giving a consistent interpretation. However, in a second experiment they gave subjects a third option and asked them to judge whether a conclusion always, sometimes, or never followed from a set of premises. In this task they found that 36 percent of the subjects gave a conditional interpretation while 15 percent gave a biconditional interpretation. It should be noted, however, that 15 percent gave responses that were consistent, but reflected neither a conditional nor biconditional interpretation. An additional 35 percent of the subjects responded inconsistently on at least one argument.

Rips and Marcus (1977) used a similar design and found that interpretation was influenced by whether the context of the syllogism invited an assumption of a causal relationship between the antecedent and the consequent. In the causal context, approximately equal numbers of subjects gave conditional and biconditional interpretations. However, in what might be described as the arbitrary context the conditional interpretation was favored by approximately three to one. Once again however, it must be noted that only 43 percent of the responses could be classified as representing either a conditional or biconditional interpretation.

In summary, the Taplin (1971) and the Taplin and Staudenmayer (1973) studies, as well as the Rips and Marcus (1977) study, show varied proportions of conditional and biconditional interpretations. However, all three found a

sizable percentage that did not respond in a consistent truth functional manner. On the other hand, Johnson-Laird and Tagart (1969) found most subjects responding in a manner consistent with a three-value conditional interpretation.

Marcus and Rips (1979) examined differences between these studies that might account for the difference in consistency of responding. An important procedural difference was that Johnson-Laird and Tagart (1969) had subjects evaluate whether the antecedent and consequent values presented on a card indicated that the rule was true, false, or irrelevant. In the other studies, subjects were presented with a conditional syllogism and asked to judge the logical validity of the argument.

Marcus and Rips (1979) employed both methods to compare performance in a single experiment. Their results were consistent with both sets of earlier studies. When asked to evaluate the truth of the rule, as Johnson-Laird and Tagart had done, less than 20 percent of the subjects provided a response pattern that could not be interpreted as either conditional or biconditional. When asked to judge the logical validity of syllogisms, as in the Taplin studies, they found that approximately 50 percent of the responses were logically contradictory. In both tasks, more conditional interpretations were reported than biconditional interpretations. However, the proportion varied considerably as a function of statement content.

For the present study there are two important points to be realized from these studies on the evaluation of conditional rules and conditional syllogisms. First, the vast majority of subjects can and do provide responses that are logically consistent when given the evaluation task, despite the fact that many of the same subjects respond in an inconsistent or logically contradictory manner on the syllogism task. Secondly, the type of interpretation (conditional versus biconditional) depends largely on the content of the If P, Q statement. The first point is important because it emphasizes the effect of subtle differences in the task that can influence subjects' responses. The second is important because it demonstrates that at least two logically consistent interpretations of If P, Q are available.

A couple of studies have examined individual differences in interpretation in the context of the selection task itself. Beattie and Baron (1988) studied the selection task and a multicard variant of it. Analysis of subject protocols revealed that some subjects interpreted the rule as a biconditional. In one experiment, after the subjects had made their selections, they were presented with evidence that contradicted the rule (i.e., a P, NOT-Q card). Approximately one-half of the subjects recognized that this evidence falsified the rule. The remaining subjects either admitted that the rule was false for that card while still holding to its overall truth, or failed to see any

significance for the card. To hold that the rule is still true may be a result of interpreting the If P, Q statement as an indication of a probabilistic relationship. If the rule described a probabilistic relationship, then the existence of one instance not following the rule would be of little consequence. To invalidate such a rule it would be necessary that a preponderance of the instances did not conform to the rule.

Another study that looked at individual differences in interpretation of conditionals (presented in the form "Every P has a Q" instead of If P, Q) on the selection task found that most subjects made selections that were appropriate to their interpretation if one assumed that the relationship was not reversible (Smalley, 1974). If the relationship is reversible, then if NOT-Q on the other side of the P card falsifies the rule, P on the other side of the NOT-Q card would also falsify the rule. Smalley argues that some subjects do <u>not</u> assume reversibility. For example, if the rule were conditional but not reversible, then P would be the only value in need of further investigation. NOT-Q would not be checked because, although the P card cannot have NOT-Q on the other side, the NOT-Q card can have any value on its back. Subjects were classified as giving a biconditional interpretation if they responded to a query by giving the biconditionally appropriate truth values (true for P, false for NOT-P) for hypothetical values found on the back of the Q card or if they stated the converse (Q-->P) of

the rule and endorsed it as being true. More than half the subjects were classified as having a biconditional interpretation on this basis.

Markovits (1988) has used a marble game to investigate subjects' understanding of If P, Q statements. The subject was shown a box with 5 small cups on top and 5 small cups on the bottom. Each of the cups on top had a hole in the bottom from which a tube extended to one of the cups on the bottom. The trajectory of the tubes was hidden except for the middle cup. The subject could see that the tube from the middle cup on top went to the middle cup on the bottom. The experimenter described various situations and asked subjects about their outcome or how they could have come about (e.g., "If a marble has landed in Cup 1/bottom, where could it have come from?"). The questions described all possible truth values for the antecedent and consequent of the conditional statement "If the marble is put in Cup 3/top, then it will land in Cup 3/bottom." After the initial trial, subjects were shown that Cup 2/top was connected to Cup 2/bottom. This manipulation was conducted to increase the perception of a one-to-one relationship. Subjects were again asked the set of questions. In the third trial, subjects were shown that the marble placed in Cup 5/top landed in Cup 2/bottom. This was done to breakdown the one-to-one bias that had been generated.

Markovits classified responses as conditional, biconditional, or intermediate. Intermediate responses were

any that did not match the conditional or the biconditional interpretation. The most common intermediate response was to draw only a modus ponens conclusion. Subjects who initially gave a conditional or biconditional interpretation to the task were very resistant to changing their interpretation. Only 10 percent of the subjects who initially gave a biconditional interpretation changed their interpretation to conditional on the third trial. Those subjects who initially gave an intermediate interpretation were much more likely to change their interpretation on the second and third trial. Markovits argues that subjects who had created a stable mental representation (conditional or biconditional respondents) were more resistant to change, and therefore, did not use the subsequent empirical evidence to modify their response. Markovits (1984, 1985) also did a series of studies showing that subjects respond to conditional reasoning tasks in a way more consistent with a material implication interpretation of a P-->Q relationship if the existence of other A-->Q relationships is emphasized (however, see Markovits, 1985, for individual differences in effect on subjects' responses).

This section has focused on the interpretation of If P, Q statements. There have been three different sources of evidence that If P, Q can have a variety of interpretations. First, there were the philosophical and linguistic arguments regarding the various uses of the If P, Q. These arguments showed a wide variety of uses for If P, Q and demonstrated

that subjects would paraphrase different uses in different ways. Second, explicit logical reasoning tasks were considered. There too interpretations varied as a function of statement content. In addition, consistent interpretation depended on the exact reasoning task being used. However, in rule evaluation studies the majority of subjects did identify the co-occurrence of P and NOT-Q as falsification of an If P, Q rule. Finally, in the selection task literature, two studies were outlined that looked at individual differences in interpretation. These studies produced evidence that some subjects interpreted the rule in the task biconditionally and, in the case of Beattie and Baron (1988), that some may have interpreted it as describing a probabilistic relationship. In light of these studies, it seems clear that If P, Q rules can have a variety of meanings that subjects pick among by looking at nonlogical factors such as the context and content of the rule.

Theoretical Explanations of Performance

A number of theoretical explanations have been proposed for how and why subjects respond as they do on the selection task. Some focus on explaining content effects, while others try to explain performance on the abstract task itself. The present study will focus on these theories primarily in their explanation of performance on the abstract task.

The typical selections on the abstract task are P alone and P & Q. An early explanation was that subjects were exhibiting a verification bias (Wason & Johnson-Laird, 1972). That is, they were only looking for cases that would make the rule true and were not recognizing the importance of potentially falsifying information. A competing explanation, known as matching bias, argued that subjects were simply matching the values named in the rule (Evans, 1972).

Evans and Lynch (1973) tested these two explanations by systematically manipulating the negation of the values found in the antecedent and the consequent. For example, with a negated consequent the rule might be: "If A, then not 5." The verification bias hypothesis predicts that the subjects would pick cards that showed an A or a number other than 5 since those would be the cards that would verify the rule. However, the matching bias hypothesis predicts that the subjects would pick A and 5 since those are the values that are named in the rule. Evans and Lynch's results confirmed the matching bias predictions. Numerous studies have consistently found evidence for matching bias; but several studies have found evidence of verification bias in addition to matching bias (Beattie & Baron, 1988; Yachanin & Tweney, 1982; Griggs & Cox, 1983) while Valentine (1985) found evidence of verification bias and not matching bias. Oaksford and Stenning (1992) also found verification bias rather than matching bias when the meaning of negation was

made unambiguous. In addition, Evans (1992) has found that using thematic content eliminates matching bias even when correct performance is not enhanced.

Evans (1984, 1989) has proposed a two-stage model of human reasoning. The first stage he calls a heuristic stage during which the relevance of the different instances are judged by preattentive processes. Matching bias reflects one of these preattentive processes. The second stage is described as an analytic stage which he leaves largely undefined. He argues that the selections made in the abstract selection task are the result of heuristic processing only (Evans 1984, 1989). Evans's model focuses on attention in that the selected cards are largely determined by whether they receive attention in the heuristic stage.

Another attentional explanation of selection task performance can be found in the work of Margolis (1987). Margolis argues, as Cohen (1981) has argued before him, that the selection task represents a "cognitive illusion" much like perceptual illusions. From this perspective, a cognitive illusion impugns the cognitive abilities no more than a perceptual illusion impugns the perceptual abilities. Margolis argues that there are two factors that make the selection task difficult. One of these factors Margolis labels semantic ambiguity referring to the multiple rule interpretations that have been described in an earlier section of this paper.

The second and more important difficulty arises from a factor that Margolis terms "scenario ambiguity." He defines two types of scenarios--open and closed. An open scenario is a situation in which one is determining how to explore different possibilities. A closed scenario occurs when the decision on how to proceed has been foreclosed, and one simply has to select from the possibilities as presented. Normally, closed scenarios are encountered only after already proceeding through the open scenario stage. The problem with the selection task is that it is a closed scenario, but may be interpreted as an open scenario.

Normally, the context is used to determine the type of problem being faced. However, the abstract selection task strips away the context and presents the problem in an impoverished environment. The subject is left to provide the context and discover the scenario. Without the appropriate semantic and scenario cues, the subject is not likely to recognize the scenario as a closed one. Margolis presents an open scenario problem, similar in structure to the selection task, that can lead to the two most common error patterns. He also proposes changes in the abstract selection task that he argues will eliminate the scenario and semantic ambiguity.

Griggs (1989) tested Margolis's claims by administering this remedial selection task in an experimental setting. Griggs found excellent performance when the scenario and semantic ambiguity were removed. However, this improvement

was only found if the rule was phrased in such a way as to make matching the value in the consequent impossible. When matching was possible, matching bias seemed to predominate.

Johnson-Laird and Byrne (1991) have recently extended their mental models theory of deductive reasoning to explain performance on the selection task. They argue that subjects construct an internal model that represents the structure and content presented in the premises of an argument. In the case of the selection task, this model contains the possible states of affairs described by the rule. Subjects only consider the cards that have been explicitly represented in their model. In evaluating their model, they will only select those cards which could have an effect on the rule's truth or falsity.

In the standard abstract problem, subjects only explicitly represent the values mentioned in the rule (P and Q). If they interpret the rule as a conditional, they will select P alone because this is the only explicitly represented card that bears on the truth of the rule. If they interpret the rule as a biconditional, then they will choose P and Q since both cards have a bearing on the rule. Thus, the mental models approach can give an account of the two most common responses on the abstract selection task.

The mental models approach also gives an explanation of correct performance. For correct performance to occur, subjects must "flesh out" their model to include NOT-Q. P must be exhaustively paired with Q, leaving NOT-Q paired

only with NOT-P. It may also be necessary to explicitly represent the impossibility of pairing P with NOT-Q. When NOT-Q is included in the model, subjects will consider it and realize its importance in evaluating the truth or falsity of the rule. Johnson-Laird and Byrne (1991) claim that subjects use information that they gain from the form, content, and context of the rule to guide the formation of the model. While Johnson-Laird and Byrne do not identify the specific mechanism that leads to the fleshing out of the model, they claim that model theory can explain all of the experimental manipulations that have shown improved performance. There are indications that the primary mechanism may be an attentional one. Anything that increases the salience of the cards usually excluded from the model may contribute to fleshing out the model.

Cheng and Holyoak (1985) have argued that people reason using pragmatic reasoning schemas. These schemas are composed of rules that outline the different relationships that are possible for that schema. The schemas are context sensitive and their use can be elicited by the content of a problem. The schema will then guide the card selections. One schema that Cheng and Holyoak (1985) have proposed is the permission schema. The permission schema links the taking of some action to the satisfaction of a precondition. Cheng and Holyoak presented a series of four production rules which make up this schema. These rules happen to lead to the same card selections that the application of the

rules of logic would produce. While Cheng and Holyoak's pragmatic reasoning schemas have been predominantly used to explain performance on concrete thematic content problems, they do introduce an abstract permission schema problem (Cheng & Holyoak, 1985, Experiment 2). Their abstract permission problem links taking action A with satisfying precondition P, where A and P remain undefined. They found over 50 percent of their subjects provided the answer predicted by the permission schema.

However, the abstract permission problem is not equivalent to the standard abstract problem. The rule in the standard abstract problem relates two general categories (letters and numbers) in an <u>arbitrary</u> relationship. The rule in the abstract permission problem relates two general categories (preconditions and actions), but the relationship is not arbitrary. The relationship between preconditions and actions is well-defined. In fact, it may be impossible to create a permission problem that is abstract in the same sense as the standard abstract problem, since the permission defines the relationship, however general, as a nonarbitrary one.

In addition, Jackson and Griggs (1990) have demonstrated that performance was not enhanced by the permission context alone, but by the presence of explicit negatives on the NOT-P and NOT-Q cards. When the explicit negatives were removed, performance was no better on the abstract permission task than on a standard abstract task.

Thus, the pragmatic reasoning schema explanation does not adequately explain the improved performance on the abstract permission task or performance on the standard abstract task.

Hoch and Tschirgi (1983, 1985) have argued that the thematic problems that provide facilitation do so because they give additional cues that the subject uses to understand the implication relationship and the basic nature of the problem. They see matching as a strategy of last resort when subjects really do not understand the relationship or the task. They report a series of studies in which logical performance on an abstract selection task was enhanced among master's-level subjects when the logical relationship represented by the rule was made clear with either implicit or explicit cues. These findings fit nicely with our previous discussion of ambiguities in the interpretation of If P, Q statements. If interpretational ambiguity is part of the difficulty in the selection task, then clarification should be expected to improve performance.

The present study hypothesizes that the difficulty of the abstract selection task originates from two sources. The first and primary source of difficulty is ambiguity in rule interpretation--what Margolis refers to as semantic ambiguity. The second source of difficulty arises from lack of analytic processing. This view fits within either of the two major theoretical frameworks. These two different

sources of difficulty can be incorporated into Evans's two-stage theory. It is during the heuristic stage that interpretation of the rule and the task must occur. Thus rule ambiguity influences the heuristic stage. If analytic processing does occur, it would be at the second stage in Evans's model. Johnson-Laird and Byrne's mental models theory can also be made amenable to the present view with model formation taking the place of the heuristic stage and model evaluation taking the place of the analytical stage.

By the present account, the common errors found on the abstract task, P alone and P & Q, result when the subject is presented with an ambiguous rule in an impoverished experimental environment where few cues exist to help disambiguate the situation. When faced with this type of task the subject opts for a matching strategy, matching the values mentioned in the rule. If further analytic processing occurs, some subjects may recognize the irrelevance of the value on the back of the Q card and will select P alone. If no further processing occurs, subjects will select P & Q. However, even when the situation is disambiguated, subjects may still adopt matching as a short-cut strategy. Matching simplifies the task and reduces the amount of analytical processing that is necessary. Therefore, as long as the task is presented in such a way that matching is possible, some use of this strategy should be expected. This view of matching differs somewhat from Evans's view. Evans sees matching as a preattentive

response bias. The present view sees matching as a processing strategy rather than a preattentive bias.

If the central source of difficulty in the task is rule ambiguity, then some improvement in performance should be expected if the nature of the rule is made more explicit. Hoch and Tschirgi have demonstrated that cues that better specify the exact nature of the relationship described by the rule enhance subjects' ability to reason about it.

The present account also predicts that performance should be improved by experimental manipulations that increase the amount of analytic processing. To this end, removal of short-cut strategies is important. In addition, the format of the task and the instructions the subjects are given about the nature of task can either enhance or detract from the analytic processing that occurs.

Thematic problems may produce facilitation because they provide the subject with more information regarding the type of relationship in question. This view is consistent with schema accounts of performance on thematic tasks. The thematic context may be triggering schemas that direct the reasoning process. In this way, schemas may provide another short-cut strategy for solving the problem.

The advantage of looking at the abstract selection task is that it provides a setting in which abstract logical knowledge and analytical ability can be examined. If the rule is made explicit but abstract, and the usefulness and salience of short-cut strategies such as matching are

systematically varied, then the subjects' ability to understand and analyze a logical relationship can be observed. The present study was designed to follow this approach.

The Abstract Selection Task

As mentioned earlier, the selection task was first described in print by Wason in 1966. Wason and Johnson-Laird (1972) have reviewed the early work on the abstract task. One of the first attempts to understand performance is a set of studies that have been described as therapy experiments. These studies were designed to facilitate subjects' performance by exposing them to information that should increase the salience of cards that could falsify the rule.

The Therapy Studies

Wason (1968) presented subjects with an abstract selection task in which the experimenter pointed to each card and asked whether knowing what was on the other side would enable the subject to determine whether the rule was true or false. In the experimental condition, the experimenter then went back through the four cards and asked the subject to indicate whether there was a value that would make the rule false if it were found on the other side of each card. Subjects were then allowed to revise their earlier selections. He hypothesized that subjects who had projected P on the other side of NOT-Q would be more likely

to see the importance of the NOT-Q card and would revise their selections in that direction.

Wason's experiment provided little support for this hypothesis. There were slightly more NOT-Q selections, but only 4 of the 18 experimental subjects selected the logically correct P & NOT-Q combination. However, this is not too surprising in view of the fact that only three subjects projected falsity in a manner that was consistent with a material implication interpretation of the rule. Of these three subjects, two revised their selections to the logically correct choice.

In a second experiment, prior to performing the selection task, subjects were asked to identify "the one card which makes the rule false." All the subjects selected the P & NOT-Q card as the falsifying case. However, this choice had no effect on subsequent performance on the selection task. The pattern of responses for subjects who had to identify the falsifying case was almost identical to that of control subjects who had no previous exposure to the P & NOT-Q card.

In a subsequent study, Wason (1969) was more successful in eliciting logically correct responses. By presenting subjects with cards that falsified the rule, first hypothetically, and then concretely, Wason was ultimately able to get almost all subjects to select the NOT-Q card. Subjects made their initial selections and then went through a series of exposures to P & NOT-Q combinations. Only two

of the 32 subjects included NOT-Q in their initial selections. After each exposure subjects were given an opportunity to revise their selections.

The first exposure was designed to present a weak hypothetical contradiction for those subjects who had not selected NOT-Q. The experimenter asked the subject what values could be on the other side of the P card. Then the experimenter asked what influence the hypothetical values would have on the truth or falsity of the rule. The experimenter continued to probe until the subject proposed NOT-Q as a hypothetical value on the other side of the P card. If the subject did not correctly identify this combination as falsifying the rule, the experimenter corrected them. Ten of the 32 subjects included NOT-Q in their revised selections following this weak hypothetical exposure.

The second hypothetical exposure was stronger. It was modeled on the first exposure, but the questions were asked about the NOT-Q card. The experimenter pointed to the NOT-Q card and asked what hypothetical values might appear on the other side. Probing continued until the subject proposed P as a hypothetical value on the other side of the NOT-Q card. If the subject did not identify this combination as a falsifying instance, they were corrected. In the revised selections following this exposure, half of the subjects included the NOT-Q card in their selections.

In the final concrete exposure, the subjects were told to turn over the cards they had selected and to see what was on the other side. All the cards had values that would conform to the rule except the NOT-Q card. If the NOT-Q card had not been among the subjects selections, the experimenter turned it over and asked the subjects if they still thought the rule was true. In the revised selections following this exposure, 30 of the 32 subjects included NOT-Q in their selections. However, even here only 10 subjects selected the logically correct combination.

In the final phase of this experiment, those subjects who had not made the correct selection were told that their selection was wrong and were asked to make a new selection. Even then, only 15 of the 32 subjects made the logically correct selection. Thus, although almost all the subjects eventually selected the NOT-Q card, less than half made the logically correct selection, even when they were told that they were wrong and were given a chance to revise their selections.

In a subsequent study, Johnson-Laird and Wason (1970b) reported similar findings. Subjects made their initial selections and then the researchers turned over the P card to reveal Q and the NOT-Q card to reveal P. Only 2 of 36 subjects picked P and NOT-Q on their initial selection, and only 9 of the remaining 34 corrected their selections immediately after seeing the other side of the NOT-Q card. Ultimately, only 23 of the original 36 subjects made the

correct selections, even after an interview in which the experimenter had the subject evaluate the information gained from each of the exposed cards.

Wason and Shapiro (1971) demonstrated somewhat improved performance when subjects first went through a familiarization deck of cards viewing one side of each card. They were told to imagine a value on the other side of each card that would make the rule true and a value that would make the rule false. The researchers corrected the subjects if they did not do this part of the task using a material implication interpretation of the rule. They found that five out of 12 subjects made the logically correct selections on a subsequent selection task. In contrast, if subjects were allowed to look at both sides of each card in the familiarization deck and simply asked to evaluate its truth value, facilitation on the subsequent selection task was not observed.

In one of the more successful therapy experiments, Legrenzi (1971) found that almost all subjects made the logically correct selections when they had previously discovered the rule from exposure to positive and negative instances of the rule. During the discovery phase, experimental subjects were told that their task was to determine the rule that governed the material found on a set of cards. They were shown the cards one at a time and they were told whether the rule was true or false for each card they saw. When they had formulated the rule, it was used in

a selection task format with four new cards. Twenty-six of the 30 experimental subjects made the logically correct selections. When control subjects were given the same rule that the discovery subjects had reported, they were much less likely to select correctly.

Another successful therapy experiment that is often overlooked was conducted by Mosconi and D'Urso (1974). They presented subjects with three sets of four cards. With the first two sets of cards, subjects were instructed to indicate "which of these cards if uncovered, could demonstrate that this hypothesis is true." For both of these so-called verification trials, 11 of the 12 subjects selected P alone. After the selection was made for each trial, the indicated cards were revealed. In these two trials, cards that violated the rule were not used.

With the last set of cards, subjects were given falsification instructions. They were told that they now had a new task which was to indicate "which of these cards, if uncovered, might demonstrate that this hypothesis is false." On the falsification trial, 8 out of 12 subjects selected P and NOT-Q. Only 1 out of 12 control subjects given the same falsification instructions but having had no previous verification trials made the P and NOT-Q selection. It should be noted that the researchers indicate that an unspecified number of subjects converted the hypothesis. They do not elaborate on what type of rule conversions took place or how they knew that the subject had converted the

rule. However, they do indicate that an unspecified correction procedure was used. Subjects who continued to convert the rule after the correction were not included. Therefore, it seems from the scant information available, that the subjects used in the study either produced a material implication interpretation or had a material implication interpretation provided.

Mosconi and D'Urso argue that the effect of the verification trials occurs because normally, falsification only becomes a strategy after a rule is considered plausible and established. In the experimental group, the exposure to the verification trials made the rule plausible. Therefore, they claim that subjects in the experimental condition are able to use the falsification strategy. Subjects in the control condition have no evidence that the rule is plausible and so they first search for verification. Thus, they predominantly select P alone or P & Q. However, these results lend themselves to an alternative interpretation.

The initial exposure to the rule in the verification task leads subjects to select P alone. When they are exposed to the falsification task the task is changed and therefore subjects are inclined to change their response. The Q card has already been rejected; and with the heavy emphasis on falsification, the subject will be inclined to introduce the NOT-Q card into their choices.

These therapy studies make several important points about performance on the selection task. The more

successful therapies provide the treatment prior to the making of an initial selection (Legrenzi, 1971) or introduce a new type of selection (Mosconi & D'Urso, 1974). That is, trying to prevent error seems more successful than trying to correct it. All of the successful therapies have in some way illuminated the nature of the rule and the role of falsifying information. The lack of facilitation in Wason's (1968) study is not surprising from this perspective. The treatment (projection of falsifying instances) did not produce responses consistent with material implication and, therefore, it is not surprising that few subjects moved to the logically correct selection. Finally, the importance of the subject actually using and actively processing the rule is illustrated in the Wason and Shapiro (1971) and Legrenzi (1971) studies. In the Wason and Shapiro study, subjects who had to actively imagine potential falsifying and verifying instances did better than those who evaluated actual cards. In the Legrenzi study, subjects had to discover the rule for themselves, presumably indicating a greater understanding of the nature of the relationship.

The therapy approach to the selection task essentially ended in 1971 with the Wason and Shapiro paper. In that paper they also reported facilitation for a problem involving concrete content. This changed the focus of much of the selection task research and the search for content effects was on. The abstract problem in most of these studies tagged along as a baseline condition to use as a

comparison for facilitation due to content. However, three other approaches to understanding the abstract task were also being pursued.

The first began in 1970 when Johnson-Laird and Wason introduced a simpler version of the selection task called the reduced array selection task (RAST). Successful performance on the RAST, along with the facilitation found for some concrete thematic problems, led to a second approach which looked for transfer from these easier problems to the more difficult abstract problem. Finally, two sets of researchers have used abstract rules that make explicit the structure of the logical relationship being described in the rule.

<u>The Reduced Array Selection Task</u>

In the reduced array selection task, subjects select between cards representing Q and NOT-Q (P and NOT-P cards are not used). Johnson-Laird and Wason (1970b) presented subjects with two stacks of cards--one composed entirely of Q cards and the other composed entirely of NOT-Q cards. Subjects were given a rule and instructed to select and inspect cards from either stack in order to prove the rule true (false). They found that all subjects exhausted the NOT-Q stack (15 cards in all). However, most selected at least some cards from the Q stack. All subjects selected at least one Q-card when they were instructed to prove the rule true. When instructed to prove the rule false, only two of seven subjects selected any Q cards. Overall, subjects

tended to be much more willing to consider NOT-Q cards when presented in the reduced array and when asked to falsify the rule.

Lunzer, Harrison, and Davey (1972) compared performance on a reduced array and a complete array selection task. In both cases, the subject was selecting from four cards. For the reduced array task, the four cards were made up of two Q cards and two NOT-Q cards. They found increased correct performance (selection of the NOT-Q cards only) for the reduced array problem when concrete content was used. However, when the content referred to shapes and colors, there was no difference in the rate of correct responses between the reduced and the complete array problems. Although there were increased selections of NOT-Q, for most subjects it was selected along with Q and not instead of it. Roth (1979) reported a similar finding even when the subjects were given a training procedure designed to clarify the conditional nature of the rule. NOT-Q selections were increased for the reduced array; but they were in addition to, not replacing, the selection of Q.

In summary, all three of these studies found that the reduced array leads to an increase in NOT-Q selections. However, Q selections remain common. Therefore, it seems clear that the difficulty of the standard selection task does not lie <u>solely</u> in a lack of attention to and consideration of the NOT-Q card.

Both Lunzer, Harrison, and Davey (1972) and Roth (1979) gave subjects pairs of problems that included both complete and reduced arrays. In this manner, they were able to look for transfer effects. The next section will consider a number of studies that examine transfer from problems that have shown improved performance like the RAST and some thematic versions of the selection task.

Transfer

The Lunzer, Harrison, and Davey (1972) reduced array study produced only limited evidence of transfer. The only case that found clear and strong evidence for transfer was in their second experiment. However, it is not transfer from the reduced array to the complete array. Instead, it is transfer from one version of the complete array to another very similar version. This transfer occurred after the subject had been given an explanation of the correct response to the first problem. Overall, subjects did slightly better on the second problem; but no clear or substantial transfer was reported.

Roth (1979) found transfer effects from the reduced array to the standard task. Initially, subjects completed either a reduced array or a complete array selection task. As described earlier the reduced array made the inclusion of NOT-Q in the selection more likely. Both groups were then given a second selection task with a complete array of cards. This increased rate of NOT-Q selections for the reduced array group transferred to the complete array task

as evidenced by a significantly higher number of P, Q, and NOT-Q selections. When on a final selection subjects were told to pick only two cards, the logically correct selection (P and NOT-Q) became the dominant selection (16 out of 28 subjects). As mentioned previously, subjects had a training session designed to clarify the logical interpretation of the rule. Additionally, subjects were asked for reasons for their selections and then allowed to revise these selections. Both of these factors may have contributed to improved performance.

Several studies have looked at transfer from thematic problems to the standard abstract task. Cox and Griggs (1982) found facilitation of logically correct performance for a thematic problem known as the drinking age problem (DAP). The rule in the drinking age problem is: "If a person is drinking beer, then the person must be over 19." The subject is told that the cards represent four people sitting around a table. The cards reveal the person's age on one side and what they are drinking on the other. They found that most subjects will make the logically correct selections given this thematic content. They found no evidence of transfer from this thematic problem to the standard abstract problem (although there was some transfer to a problem that arbitrarily linked the clothing color with age).

Berry (1983) examined transfer in a complex experiment that examined the effects of verbalization instructions

(requiring subjects to report reasons for their choices) and problem content (concrete versus abstract) for rules where the negation of the antecedent and the consequent was systematically varied. Initial performance was poor for both concrete and abstract contents.

Following the first trial, subjects were given an explanation of the correct answer. Subjects in the concrete condition showed considerable improvement on subsequent trials; but subjects in the abstract condition showed very little improvement. Although subjects in the concrete condition did well on subsequent concrete tasks, they were unable to transfer this ability to a series of abstract tasks. However, transfer to the abstract problems was substantially increased when subjects were instructed to give verbalizations or reasons for the choices that they made on the concrete task. Therefore, transfer from concrete tasks to abstract tasks did not seem to occur unless the subject was given an explanation of the correct answer on the concrete problems and was required to verbalize their reasoning process. Without the verbalization instructions, performance remained low even when subjects were told that the abstract problems were "logically identical" to the concrete problems and that they could solve them "in the same way."

Klaczynski, Gelfand, and Reese (1989) failed to replicate Berry's (1983) finding of a facilitatory effect of verbalization on transfer. They found no transfer from

thematic problems to abstract problems even when the subjects were given an explanation of the correct responses to the initial thematic problems. However, they did find transfer between abstract problems when an explanation of the correct response was given for the first problem. In their final experiment they found that transfer was greatest when the initial problem was abstract rather than thematic, regardless of whether the final problem was abstract or thematic. The largest transfer effect was observed when both initial and final problems were abstract. Little transfer was observed when the initial problem was thematic and the final problem was abstract. They concluded that the facilitatory effect of explaining the correct response to the abstract problem may be attributed to the development of a general problem-solving strategy which can be used in either abstract or thematic problems.

Lawson (1987) has also found transfer effects between abstract problems when an explanation of correct responses is given after the initial problem. He found a difference in likelihood of transfer as a function of developmental level. High school students that he classified as formal operational based on an earlier task were more likely to give the logically correct response to the posttest selection task than students that he had classified as concrete operational. He also found that performance remained high even with a one month delay between the initial exposure and the final posttest.

In summary, several experiments have shown transfer, but only when the initial problem is an abstract problem and an explanation of correct performance is provided. There is little evidence of transfer from thematic to abstract problems. In addition, initially presenting the RAST increased the likelihood of selecting the NOT-Q card in subsequent complete array selection tasks. However, generally the Q card was also included unless additional measures were taken to make the subject pick between them.

Rule Explication

Two pairs of researchers have used rules that explicitly clarify the logical relationship under consideration. The first pair, Mosconi and D'Urso (1974), were not expressly concerned with rule ambiguity. Instead they had noted that subjects seemed to disregard the NOT-Q card. They felt that if they included mention of cards without Q in the rule that this might increase the salience of the NOT-Q card. They used a rule which read: "If to the left there is the letter A, then to the right there is the number 2, and never a number different from 2." Since a "number different from 2" was mentioned they argued that this would increase the salience of the NOT-Q card. Six of 12 subjects in this condition made the logically correct selection. However, it should be noted that the instructions used were different from the true-false instructions that are typically used. These instructions emphasize looking for cards that "could demonstrate that

this hypothesis is false" or that "could be contrary to the hypothesis." This issue will be examined in more detail in an upcoming section on instructions. Additionally, some unspecified procedure was used for detecting and correcting subjects who converted the rule. The number and type of conversions is not specified, and "obstinate converters" were not included in the data. Despite these two cautions, this experiment does lend some support to the idea that elimination of rule ambiguity may help facilitate performance.

A second set of studies have already been briefly mentioned. Hoch and Tschirgi (1983, 1985) report a series of experiments in which they give implicit and explicit cues to the logical relationship under consideration. Subjects were also required to give a reason for the decision they made about each card. In the explicit condition, a clarification statement was added following the usual If P, Q rule. The clarification told subjects: "Cards with the letter A on the front may only have the number 2 on the back, but cards with the letter B on the front may have either 1 or 2 on the back." This clarification improved performance although it was mediated by a level of education effect (see Hoch & Tschirgi, 1985). In addition, the analysis of the subjects' reasons revealed evidence that subjects in the subjects provided with rule clarification considered more hypothetical card combinations in their analysis.

Both pairs of researchers provide evidence that subject performance can be enhanced with rule explication. While the Mosconi and D'Urso (1974) study used an unusual instruction format, Hoch and Tschirgi (1985) used standard true-false instructions and still found facilitation, at least at higher education levels. However, Hoch and Tschirgi (1983, 1985) also introduced an unusual procedural format. They required subjects to provide written reasons explaining why each card was or was not selected. The next section reviews several of these format and instructional variables that may influence performance on the abstract selection task.

Format and Instructional Factors

Violation Instructions

The early abstract selection task problems asked the subject to select the cards necessary to determine whether the rule is true or false. When thematic problems such as the drinking age problem were introduced, the instructions told the subject to select the cards that need to be turned over in order to determine whether the rule has been violated. The effect of instruction type (true-false versus violation) has been examined in both thematic and abstract problems.

Several researchers have found that violation instructions increase facilitation for thematic problems (Chrostowski & Griggs, 1985; Dominowski, 1989; Griggs, 1984; Yachanin, 1986). However, for abstract problems

facilitation has been elusive. Chrostowski and Griggs (1985), Griggs (1984), and Yachanin (1986) all found no improvement in the abstract selection task when presented in a violation format. Mosconi and D'Urso (1974) used instructions that emphasized falsification. While these were not the typical violation instructions, they resembled violation instructions more closely than true-false instructions. The instructions asked the subject to identify cards that could make the rule "false" or that were "contrary to" the rule. In an explicit rule condition, 6 of the 12 subjects made the logically correct selection. Although only 1 of 12 subjects produced the logically correct response with the same instructions and a standard If P, Q rule, the instructions could have interacted with the type of rule contributing to the 50 percent correct response rate in the explicit rule condition. Jackson and Griggs (1990, Experiment 1) found improvement with violation instructions in a standard abstract selection task, but the difference was not replicated in a subsequent experiment. Thus, while increased facilitation produced by the use of violation instructions seems clear for thematic problems, there is little evidence of its effectiveness in the abstract problem.

<u>Reason-Giving</u>

Several studies have reported that subjects asked to give reasons for their choices, or to reason aloud while they solved the problem, show better performance on abstract

problems. Hoch and Tschirgi (1985) and Jackson and Griggs (1988) had subjects give reasons for their initial choices and then allowed subjects to indicate a second choice if they felt their initial choice was incorrect. Roth (1979), in a transfer study described earlier, gave subjects an opportunity to revise their selections after giving reasons for their choices. Although all three sets of researchers reported a higher number of logically correct selections for the second-choice data, the increases were not statistically significant.

Berry (1983) used a complex design, described earlier, that did not test reason-giving independent of other procedural changes. However, some evidence of a possible reason-giving effect was reported. She had some subjects verbalize their reasoning as they attempted to solve a concrete selection task problem. Performance on this initial problem was poor. After the task, subjects were given an explanation of the correct answer and then given additional trials using concrete problems. On these subsequent problems, performance exceeded 80 percent correct for all concrete conditions. There was no difference in performance between verbalization and no verbalization conditions. However, subjects in the verbalization condition did significantly better than nonverbalizers on a subsequent abstract transfer task.

Because of the explanation of correct performance and the multiple-problem format, interpretation of the reason-

giving effect found in Berry's (1983) study is difficult. Chrostowski and Griggs (1985) conducted a study using a similar verbalization procedure, but with a single-problem format and without the explanation of correct performance. They failed to find an effect of verbalization instructions for subjects solving individual problems. However, they also reported difficulty in getting the subjects to comply with the verbalization instruction.

Yachanin (1986) required subjects to explain why each card did or did not need to be turned over for both concrete and abstract problems in a study of the effect of instructions (true-false versus violation). All subjects were instructed to give reasons so we are unable to look at the effect of this manipulation systematically. However, logically correct performance was very high (100 percent for one problem) for the concrete problems when violation instructions were included. Despite this finding, performance on the abstract problem was poor even with violation and reason-giving instructions.

Dominowski (1990) reported a study that looked at the effect of giving reasons on a wide variety of selection task problems. He overcame the problem of noncompliance with the verbalization instructions by having subjects write down their reasons. Overall, those subjects giving reasons did better; but the difference was small for the problems that would be considered abstract. This effect was also accompanied by a procedural variation in which subjects had

to make a two-alternative forced-choice on each card by checking spaces marked either "yes" or "no."

Van Duyne (1976) asked subjects to produce rules which were either necessary (always true) or contingent (sometimes true) conditional relationships. The experimenter selected two sentences to be used in a selection task. Subjects were told to imagine that the experimenter was unfamiliar with the rule and wanted to assess whether it was true or false. Subjects were given each logical case (P, Q, NOT-P, and NOT-Q) and asked if the experimenter should seek more information for that instance. Subjects were asked to give a reason for each decision.

Van Duyne used these reasons along with the actual responses to classify whether subjects had achieved complete logical insight in solving the problem. A selection was classified as insightful only if the reason given for the selection was a logical one. He found that NOT-Q selections were quite common and that 7 out of 22 subjects achieved complete insight in the contingency condition. In the necessity condition, complete insight was less common; but the selection of NOT-Q remained frequent.

Interpreting the reasons given as reflecting the actual processes underlying the selection decision has been criticized. Evans and Wason (1976) had subjects give reasons for solutions that the experimenters provided and found evidence that the reasons were simply post hoc rationalizations. However, reasons that are given

concurrent with the selection process seem more likely to represent actual analytic processing by the subject. The increase in correct selections when asked to provide reasons is evidence that, in at least some cases, the reasons may reflect analytic processing. Asking subjects to provide reasons for their choices may force them to use analytic processing in cases where they might otherwise have used a short-cut strategy such as matching.

The effect of reason-giving in the experiments that have been described in this section is far from clear. A number of studies found no significant effect of reason-giving on the abstract selection task. When significant results were found, they were accompanied by other procedural changes such as the explanation of correct choices used by Berry (1983) and the forced-choice response format used by Dominowski (1990). A systematic examination of these variables is necessary to accurately assess the effect of giving reasons on selection task performance.

Forced-Choice

The forced-choice response format used by Dominowski (1990) has also been used by other researchers but has never been systematically manipulated. Forcing a choice on each card could contribute to facilitation since it forces the subject to attend to each card at least briefly. Yachanin (1986, Experiment 1) used a forced-choice format which may have contributed to the 100 percent correct response rate that he found for the drinking age problem when he combined

violation instructions and reason-giving for each individual card. However, he did not find a similar effect for an abstract version of the task (Experiment 2). Additionally, Yachanin and Tweney (1982) used a forced-choice format but found typical poor performance for an abstract problem and three concrete problems. However, the concrete problems used were not ones that have reliably produced correct responding.

Pollard (1981) used a forced-choice procedure with an abstract selection task and a concrete selection task. He found slight facilitation for the concrete version, but typical poor performance for the abstract version. He attributes the facilitation to the concrete content, but the particular type of content used, pairing particular towns with particular modes of transportation, has not produced reliable performance (for a review see Griggs, 1983). It is also possible that the forced-choice format contributed to the facilitation in performance observed on the concrete selection task.

Pollard and Evans (Experiment 2, 1987) found higher than usual levels of performance for abstract selection task problems using the forced-choice format. However, they also changed the rule so that the consequent was "there is a number greater than 18 on it." This does not allow the subject to match the topic of the consequent and so could change the way the subject responds. In Experiment 1, they used a more typical letter/number rule along with the

forced-choice format. None of the subjects gave the logically correct response with this more typical abstract rule.

Cheng and Holyoak (1989) report an experiment in which they introduce precaution problems. Their abstract precaution rule links engaging in hazardous activity H with having protection P. They also introduced thematic precautions which replaced H and P with actual activities and protections. These problems fit either permission or obligation schemas, depending on whether the activity is voluntary for the actor. Cheng and Holyoak (1989) found high rates of correct performance for the abstract and thematic precaution problems. They attribute this result to the activation of pragmatic schemas. However, the interpretation of this result is confounded because they used the forced-choice format (Holyoak, personal communication). They also report a relatively high baseline for their nonprecaution problem (28 percent correct). This lends further evidence that forced-choice may have contributed to the facilitation observed.

Although the use of a forced-choice format has not been systematically investigated, it bears consideration as a potential facilitating factor. None of the studies described have shown higher than usual performance for the standard abstract problem when matching was possible. However, several of the studies do indicate a potential increase in facilitation for thematic problems that are

already known to produce some facilitation. This increase may be attributable to the forced-choice format.

This section has reviewed three different instructional and format variable and their influence on the abstract selection task. In general, these variables had very little effect on the abstract selection task. Griggs and Cox (1992) considered the effect of some other instructional and format variables and their influence on both Cheng and Holyoak's (1985) abstract permission problem and a standard abstract problem. These variables were manipulated in two factorial experiments. The presence of explicit negatives on the NOT-P and the NOT-Q cards and the use of a violation instruction variant that restricted selections to two cards enhanced performance in the abstract permission problem. Correct performance in the abstract permission problem varied from 5 to 85 percent across conditions in the two experiments. However, correct performance did not exceed 10 percent in all but one of the conditions using the abstract selection task. In the remaining condition, correct performance was only 25 percent. Thus, despite the large impact of these two variables on performance in the abstract permission problem, there was no significant effect on correct performance in the standard abstract problem.

General Overview of Study

The introduction has described two major sources of selection task difficulty. First, the If P, Q rule can have multiple interpretations, and therefore, its meaning may be

ambiguous. Many of the failed attempts at facilitation in the abstract selection task literature may have been a result of rule misinterpretation. Those studies that have produced facilitation in the abstract selection task have implicitly or explicitly removed the rule ambiguity. Secondly, a lack of analytic processing can lead to poor performance even when the rule ambiguity is eliminated. If strategies such as matching are viable, subjects may opt for the short-cut and by-pass more careful analysis. In addition, instructional and format variables may influence the amount and type of processing. While most of the studies of these variables have found no effects in the abstract task, this may be due to the presence of rule ambiguity. These instructional and format variables need to be investigated in the context of a explicit rule.

The present study investigated the rule ambiguity/lack of analytic processing hypothesis. Following the lead of Hoch and Tschirgi (1983), Experiment 1 made explicit the nature of the relationship described by the rule. All four card values were mentioned in the rule so matching should be reduced or eliminated as a short-cut strategy. If the primary obstacle to logical performance on the abstract selection task is rule ambiguity, then the explicit conditions in Experiment 1 should show improved performance.

Experiment 2 compared performance on the standard rule and one of the explicit rules introduced in Experiment 1. The use of a reason-giving response format was introduced in

a factorial design. Reason-giving was combined with a forced-choice format so that subjects would have to consider and reason about each card. It was hypothesized that this reasons response format would increase correct responding by discouraging the use of short-cut strategies and increasing analytic processing. However, this effect should only be seen when the nature of the rule has been made explicit.

Experiment 3 introduced an explicit rule that allows for the use of a matching strategy. All four logical conditions were mentioned in the explicit rules used in Experiments 1 and 2. In Experiment 3, only P and Q were mentioned in the explicit matching rule. This makes matching a viable strategy for differentiating between the cards. The hypothesis was that when matching can be used to differentiate between the cards, correct performance will be reduced due to a reduction in analytic processing. Since the nature of the conditional relationship was made explicit, an additional hypothesis was that the matching explicit rule would increase the selections of P alone since the subject would not interpret the rule as a biconditional. When the reasons response format was used, increased correct performance for both types of explicit rules was expected since analytic processing would be required. However, the improvement should be least for the matching rule since more complex processing can still be avoided by using the matching strategy.

True-false selection instructions were used in Experiments 1, 2, and 3. In Experiment 4, the effects of true-false and violation instructions were examined. Because the violation instructions should facilitate processing of the NOT-Q card, the hypothesis was that the violation instructions would produce superior performance.

CHAPTER 2
EXPERIMENT 1

Experiment 1 examined the role of rule ambiguity in the production of the typically poor performance found in the standard abstract selection task. Results for the standard abstract problem were compared to results for conditions that had an explication statement explaining the logical relationship described by the standard If P, Q rule. If rule ambiguity plays a major role in creating the difficulty of the abstract selection task, then the explication should improve performance.

Hoch and Tschirgi (1985) presented subjects with an abstract selection task containing an explication statement. Although, they found 72 percent of subjects with a Masters degree made the correct selection, only 24 percent of subjects with a high school degree (comparable in education to the undergraduates used in the present study) made the correct selection. While a 24 percent correct response rate is better than what is typically found for the standard abstract problem without explication, it is not a large improvement. In Hoch and Tschirgi's study, the explication statement accompanied the standard If P, Q rule. Pilot work for the present study indicated that subjects may be confused by what they perceive as two rules--the If P, Q

rule and the explication statement. Therefore, in the present Experiment 1, the explication statement replaced, rather than supplemented, the standard If P, Q rule in some conditions.

In Experiment 1, the presence of the standard rule and the presence of an explication were systematically varied. When the standard rule was absent, the explication took the place of the rule. The explications mentioned all four card values and were composed of four different wordings to look for attentional effects due to the position of each card value in the statement. For example, the card mentioned first in the explication might be seen as the focus of the explication and therefore be more likely to be selected. If this type of matching strategy was used, it could be detected by comparing individual card frequencies for the different explication wordings. However, this type of matching effect was not predicted because all four values appear in the rule.

The rule ambiguity/lack of analytic processing hypothesis would predict that correct performance would be highest in the explication conditions. In addition, if the suggestions of the pilot work are correct, the highest performance among the explication conditions would occur when the standard rule is absent. The rule ambiguity/lack of analytic processing hypothesis would also predict that the more typical matching pattern, selecting P & Q or P alone, would be highest in the standard-rule-present

conditions since the presence of the standard rule would make matching a viable short-cut strategy.

Method

Subjects

One-hundred and eighty students from introductory psychology classes at the University of Florida served as subjects to partially fulfill a course requirement.

Design

Each subject was given a single selection task problem. Subjects were run in groups of 5-10 with 20 subjects assigned to each of nine conditions. The nine conditions are the result of a partial 2 x 5 factorial design with the standard rule (present or absent) and the type of explication (none or one of the four different wordings described in the materials section) as the two factors. The standard rule absent, no explication condition is omitted since no rule statement of any type would occur in this version. Four card orders were used to control for card position effects. They were P, NOT-P, Q, NOT-Q; NOT-Q, Q, NOT-P, P; NOT-P, NOT-Q, P, Q; and Q, P, NOT-Q, NOT-P.

Materials

The basic abstract problem can be found in Appendix A, and is a standard abstract problem with true-false instructions. Four types of explications were used and are identified by the logical value referred to first in the sentence. They are:

P explication: A card with a vowel on it can only have an even number, but a card with a consonant on it can have either an even or an odd number.

NOT-P explication: A card with a consonant on it can have either an even or an odd number, but a card with a vowel on it can only have an even number.

Q explication: A card with an even number on it can have either a vowel or consonant, but a card with an odd number on it can only have a consonant.

NOT-Q explication: A card with an odd number on it can only have a consonant, but a card with an even number on it can have a vowel or a consonant.

In the standard-rule-present/explication-present conditions, the explication immediately followed the standard rule and was introduced with "This means that". In the standard-rule-absent conditions, the explication replaced the standard rule.

Results and Discussion

The percentage correct for each condition can be seen in Table 2-1. Throughout the literature a number of different approaches to analyzing selection task data have emerged. Some reflect theoretical differences, while others are simply different statistical approaches. The present study will report and analyze the data on the basis of correct response frequencies, Pollard and Evans's (1987) logic and matching indexes, and individual card frequencies.

Pollard and Evans's (1987) logic index is calculated by adding 1 point for each logically correct card choice (P or NOT-Q) and subtracting 1 point for each logically incorrect card choice (NOT-P or Q). Their matching index is calculated by adding 1 point for each matching card choice (P or Q) and subtracting 1 point for each nonmatching card choice (NOT-P or NOT-Q). The resulting indexes can vary from 2 to -2. The logically correct combination, P & NOT-Q, would result in a logic index of 2 and a matching index of 0. The matching combination, P & Q, would result in a logic index of 0, since the correct choice of P would be offset by the incorrect choice of Q, and a matching index of 2. The logic index scores and the matching index scores can be seen in Table 2-2 and 2-3, respectively.

The response pattern for the standard rule without explication exhibited the usual errors (16 of 20 selected either P alone or P & Q). One half of the subjects selected the P & Q matching response. Two-way analyses of variance were calculated on correct responses and logic index scores for the explicated conditions. Correct responses were less frequent when the standard rule was present (20 percent versus 31 percent) but this difference failed to reach significance, $F(1,152) = 2.60$, $p = .11$. With the standard rule absent, correct responses ranged from 25-40 percent. No other main effects or interactions approached significance.

Since the only source of differences in correct response rates seems to be the presence or absence of the standard rule, the explication conditions were collapsed across the four statement wordings within each level of the standard rule. These two collapsed groups were individually compared to the standard abstract problem. The standard rule absent condition showed a rate of correct responses that was significantly higher than the rate for the standard abstract task, $X^2(1, \underline{N} = 100) = 8.33$, $\underline{p} < .01$. The standard rule present condition also showed a significantly higher rate of correct performance than the standard abstract task (Fisher's Exact Test, 1-tail, $\underline{p} = .02$). Fisher's Exact Test was used because the small expected frequency in one of the cells violated the assumptions of the Chi-square test.

A two-way analysis of variance was also conducted on the matching index scores for the explication conditions. There were no significant main effects or interactions. However, the main effect of Rule Type and the main effect of Explication Type both approached significance, $\underline{F}(1,152) = 2.77$, $\underline{p} = .10$, and $\underline{F}(3,152) = 2.55$, $\underline{p} = .06$, respectively. There were higher matching index scores when the standard rule was present (0.56 versus 0.29). Higher matching index scores were also found for the P and NOT-P explication conditions, 0.67 and 0.60, respectively. In the Q and NOT-Q explication conditions, matching index scores fell to 0.10 and 0.32, respectively. The matching index score for the standard-rule-present/explication-absent condition was

significantly higher than the overall matching score for the eight explication-present conditions (1.25 versus 0.425), $t(178) = 3.74$, $p < .001$.

The proportions of individual card choices can be found in Table 2-5. Four two-way analyses of variance were conducted to examine individual card choices in the explication conditions. For these analyses the selection frequency for each individual card was analyzed independently. There was a significant main effect of Rule Type on the frequency of NOT-Q selections, $F(1,152) = 5.05$, $p < .05$. When the standard rule is present, NOT-Q selections are less frequent (.34 versus .51). There were no other significant main effects or interactions for any of the individual card choice analyses.

The analyses of individual card choices for the explicated conditions did not reveal any matching bias for the card mentioned first in the explication since there were no significant effects of Explication Type. Informal observation of the error pattern for the explication conditions revealed two different response patterns when the standard rule was absent. The P and NOT-P explications show one response pattern, while the Q and NOT-Q explications show another. For the P and the NOT-P explications, the error pattern is fairly typical with P alone and P & Q selections predominating. The only exception is the slightly higher than usual rate of selecting Q alone. However, for the Q and NOT-Q explications, the error pattern

looks different. P & Q selections are reduced and the normally rare NOT-P & NOT-Q selection increased.

If P represents vowels and Q represents even numbers, then the P and NOT-P explications describe the implication rule P implies Q, while the Q and NOT-Q explications describe the contrapositive of that rule, which is the logically equivalent, NOT-Q implies NOT-P. Therefore, the choice of P has the same logical status as the choice of NOT-Q and the choice of Q has the same logical status as the choice of NOT-P. When the standard rule is present, it can guide choices to the matching P and Q selections and away from the nonmatching NOT-P and NOT-Q selections. The absence of the standard rule may help explain the NOT-P and NOT-Q selections found in Q and NOT-Q explications.

Subjects seem to have a bias toward picking the cards mentioned in the restrictive clause of the rule. For example, for the P explication, subjects seem to match the values described in the clause "A card with a vowel on it can only have an even number." This seems to describe the error pattern for P and NOT-P explications fairly well. It also helps explain the direction and type of change in errors for the Q and NOT-Q explications. However, the Q and NOT-Q explications also show common P & Q selections. Thus, while there does seem to be some evidence of a matching bias in the response pattern, it is weaker and less reliable than the typical bias found with the standard rule.

The results of this experiment provide evidence that rule interpretation is a likely factor in the difficulty of the selection task. However, even in the best explication condition only 40 percent of the subjects made the logically correct response. There was evidence of matching bias in the standard-rule-present/explication-absent condition. Matching index scores decreased considerably when an explication was present. Some evidence of a weak matching bias was found in the explication conditions; but, in general, the results for these conditions just look noisier with no particular selection combination predominating. While subjects seem to do better when the usual matching strategy cannot be used and when the rule is explicit, obstacles still remain that prevent the majority of subjects from correctly responding to the selection task.

TABLE 2-1
Percents Correct for Experiment 1

Standard Rule	Type of Explication				
	None	P	NOT-P	Q	NOT-Q
Present	0	15	20	20	25
Absent	---	35	40	25	25

Note. "None" is the standard abstract problem without any explication. The type of explication is identified by the logical case appearing first in the sentence (e.g., P-explication: "A card with a vowel on it can only have an even number, but a card with a consonant on it can have either an even or an odd number."). _N_ = 20 for each condition.

TABLE 2-2
Logic Index Scores for Experiment 1

Standard Rule	Type of Explication				
	None	P	NOT-P	Q	NOT-Q
Present	0.35	0.05	0.50	0.10	0.60
Absent	---	0.50	0.90	0.70	0.25

TABLE 2-3
Matching Index Scores for Experiment 1

Standard Rule	Type of Explication: None	P	NOT-P	Q	NOT-Q
Present	1.25	0.65	0.70	0.50	0.40
Absent	---	0.70	0.50	-0.30	0.25

TABLE 2-4
Frequency of Selection Combinations for Experiment 1

Standard Rule Absent

	Type of Explication			
Selection	P	NOT-P	Q	NOT-Q
P	2	2	3	2
NOT-P	1	0	1	0
Q	3	2	0	4
NOT-Q	0	1	3	1
P, NOT-P	0	1	1	1
P, Q	5	4	1	2
P, NOT-Q*	7	8	5	5
NOT-P, Q	1	0	1	2
NOT-P, NOT-Q	0	1	3	2
Q, NOT-Q	0	0	1	1
P, NOT-P, Q	0	0	0	0
P, NOT-P, NOT-Q	0	0	1	0
P, Q, NOT-Q	0	1	0	0
NOT-P, Q, NOT-Q	0	0	0	0
ALL	1	0	0	0
NONE	0	0	0	0
TOTAL	20	20	20	20

*Correct answer

TABLE 2-4 continued

Standard Rule Present

	Type of Explication				
Selection	None	P	NOT-P	Q	NOT-Q
P	6	3	4	1	6
NOT-P	0	2	0	3	0
Q	0	0	4	2	3
NOT-Q	1	0	1	1	3
P, NOT-P	2	0	1	0	0
P, Q	10	7	4	5	1
P, NOT-Q*	0	3	4	4	5
NOT-P, Q	0	3	0	2	2
NOT-P, NOT-Q	0	1	1	0	0
Q, NOT-Q	1	0	0	1	0
P, NOT-P, Q	0	0	0	0	0
P, NOT-P, NOT-Q	0	0	0	0	0
P, Q, NOT-Q	0	0	1	1	0
NOT-P, Q, NOT-Q	0	0	0	0	0
ALL	0	1	0	0	0
NONE	0	0	0	0	0
TOTAL	20	20	20	20	20

*Correct answer

Table 2-5
Proportions of Individual Card Choices for Experiment 1

Standard Rule Absent

Card	Type of Explication P	NOT-P	Q	NOT-Q
P	.75	.80	.55	.45
NOT-P	.15	.10	.35	.25
Q	.50	.35	.15	.45
NOT-Q	.40	.55	.65	.45

Standard Rule Present

Selection	Type of Explication None	P	NOT-P	Q	NOT-Q
P	.90	.70	.70	.55	.60
NOT-P	.10	.35	.10	.25	.10
Q	.55	.55	.45	.55	.30
NOT-Q	.10	.25	.35	.35	.40

CHAPTER 3
EXPERIMENT 2

In Experiment 1, correct performance was enhanced by the use of an explicit rule. However, more than half of the subjects still gave logically incorrect answers. As reported earlier, Hoch and Tschirgi (1985) also found only marginal facilitation with the introduction of an explication statement for subjects with a level of education comparable to the subjects in the present Experiment 1. However, the present Experiment 1 and Hoch and Tschirgi's study differed in two important ways. First, in Hoch and Tschirgi's study, an If P, Q rule was included in the explicit condition. The present Experiment 1 found that the presence of an If P, Q rule reduced correct performance. Second, Hoch and Tschirgi required subjects to provide reasons for their decisions about each of the cards. The present Experiment 1 made no such provision. Requiring subjects to provide reasons may facilitate analytic processing. Therefore, the rule ambiguity/lack of analytic processing hypothesis would predict better performance with the use of a reasons response format. In Experiment 2, rule type (standard versus explicit) and the presence or absence of a reasons response format were systematically varied.

As reviewed earlier, several researchers have found that requiring reasons for selections increases logically correct performance. In addition, a number of researchers have used a forced-choice response format that required subjects to respond either yes or no to each card. Both of these factors should increase the analytic processing of each of the cards. However, pilot work indicated an unreliable effect of forced-choice alone. In Experiment 2, forced-choice was not included as a independent factor, but the reasons response format included forced-choice so that subjects would have to give a justification for their decision about each card, regardless of whether it is one of the cards that the subject decides to turn over. Hoch and Tschirgi (1983, 1985) used this type of response format and analyzed the reasons their subjects gave by classifying the individual card choices based on whether they explicitly reflected an inference based on a conditional rule interpretation. For example, a reason for not selecting Q that said Q could be paired with either P or NOT-P was counted as a logical inference, while one that simply said Q was irrelevant was not. They did this type of analysis in order to examine problem type effects on the hypothetical card combinations considered. For instance, in the above example, the Q paired with P or NOT-P reason implies two hypothetical combinations while the Q is irrelevant reason implies none.

Given that such a combinatorial analysis for problem types was not our objective, the reasons in the present Experiment 2 were analyzed differently. The reasons for each subject were placed in post hoc categories based on the reasons given for all four cards rather than for individual cards. By looking at the reasons given for all four cards, common interpretations of the rule, both logical and nonlogical, and how these interpretations relate to all of the cards may be revealed. This should produce a more complete picture of the subjects' perception of the rule. It also allows for detection of inconsistencies in rule interpretation that might be missed in an analysis based on individual cards.

An additional advantage of requiring reasons is that they may reveal misunderstanding about the nature of the task itself as Beattie and Baron (1988) have reported. The reasons provide the experimenter with another avenue for determining exactly what the subject perceives the task to be.

Method

Subjects

Forty-eight students from introductory psychology classes at the University of Florida served as subjects to partially fulfill a course requirement. None of the subjects had participated in Experiment 1.

Design

Subjects were run in small groups, and each subject was given one selection task problem. Twelve subjects were assigned to each of four conditions. The four conditions were the result of a 2 x 2 factorial design. The variables were: Rule Type (standard or explicit) and Reasons (present or absent). Four card orders were used to control for card position effects. They were P, NOT-P, Q, NOT-Q; NOT-Q, Q, NOT-P, P; NOT-P, NOT-Q, P, Q; and Q, P, NOT-Q, NOT-P.

Materials

In the explication conditions of Experiment 1, performance was at its lowest level when the standard rule was present. Although there was no significant difference in correct responding as a result of which explication statement was being used, the P and the NOT-P explications led to the best performance. Therefore, the explicit rule used was the P-explication without the standard rule. For the reasons-present conditions, space for reasons was provided next to each card prompted by the word "Reason:" and the statement "Give a reason, in the space provided, for the decision you made about each of the four cards" was added to the instructions. A choice was forced in the reasons-present conditions by placing "Turn over? Yes_____ No_____" next to each card. The sentence "Circle only the card or cards that must be turned over" was replaced with "Indicate which card or cards must be turned over by placing checkmarks in the appropriate spaces."

Results and Discussion

The results of Experiment 1 demonstrated improved performance for an abstract selection task with an explicit rule. However, the majority of subjects still made errors indicating that an explicit rule may be necessary but not sufficient for facilitation of performance in the abstract task. In Experiment 2, requiring reasons for subjects' choices was examined as a potential facilitating factor. In addition, requiring reasons also offers the researcher a second approach to understanding subjects' errors. Therefore, in addition to the usual quantitative analysis, a qualitative analysis is included that attempts to describe the types of reasons and what they may reveal about the subjects' interpretation of the rule and the task.

Quantitative Analysis

Percents correct and logic index scores for Experiment 2 can be found in Table 3-1. Two-way analyses of variance were performed for correct responses and logic index scores. Rule Type produced a significant main effect for both correct responses, $F(1,44) = 13.10$, $p < .001$ and logic index scores, $F(1,44) = 6.96$, $p < .05$. The explicit-rule conditions had a 50 percent correct response rate overall, while the standard-rule conditions had a correct response rate of only 8.3 percent. There was also a main effect of Reasons on correct responses, $F(1,44) = 4.71$, $p < .05$. When asked to provide reasons, 41.7 percent of subjects selected correctly, while only 16.7 percent of those not asked to

give reasons made the correct selections. There were no significant interactions.

Matching index scores were calculated and can be found in Table 3-2. A two-way analysis of variance revealed two significant main effects. First, as expected, there was a main effect of Rule Type, $F(1,44) = 7.77$, $p < .01$. Consistent with the findings of Experiment 1, the matching index for the standard rule conditions was 1.04 compared to a matching index of only 0.33 for the explicit rule conditions. Secondly, there was a main effect of Reasons, $F(1,44) = 7.77$, $p < .01$. When reasons were absent the matching index was 1.04. The matching index fell to 0.33 when the reasons were present.

The frequency of selection combinations can be seen in Table 3-3. The standard-rule conditions produced a typical error pattern with P alone and P & Q selections dominating. When reasons were absent the P & Q selection was predominate, but with the introduction of reasons, the P alone and P & Q selection frequencies were more equally split. While the error pattern was a little more varied for the explicit-rule/reasons-absent condition, P alone and P & Q selections are still most common. In the explicit-rule/reasons-present condition, errors were infrequent and idiosyncratic. The qualitative analysis of the reasons given will provide an opportunity to further investigate why some of these responses were made.

The proportion of subjects making each individual card selection can be found in Table 3-4. A two-way analysis of variance was conducted on the individual card frequency data for each of the four card. There were no significant main effects or interactions for the P selection analysis. P selections predominated in all four conditions. The NOT-P selection analysis revealed a significant interaction between Rule Type and Reasons, $F(1,44) = 4.23$, $p < .05$. In the standard-rule condition, the rate of NOT-P selections decreased when reasons were present (.08 versus .25). This can be traced to a reduction in the rate of selecting all the cards. In the explicit-rule conditions, none of the subjects selected NOT-P when reasons were absent. When reasons were present, 25 percent selected NOT-P. This increase can be traced to the presence of NOT-P selections in the idiosyncratic errors mentioned earlier.

There were main effects of both Rule Type and Reasons in the analysis of Q selections, $F(1,44) = 7.82$, $p < .01$ and $F(1,44) = 12.22$, $p < .01$, respectively. In addition, there was an interaction between Rule Type and Reasons, $F(1,44) = 4.40$, $p < .05$. The interaction is driven by a extremely high proportion of Q selections in the standard-rule/reasons-absent condition. Ninety-two percent of the subjects in this condition included Q in their selections. This dropped to 25 percent when reasons were present. The presence of reasons also reduced Q selections in the explicit rule condition, but the initial level of Q

selection when reasons were absent was only 33 percent. Finally, the analysis of NOT-Q selections revealed a significant main effect of Rule Type, $\underline{F}(1,44) = 7.62$, $\underline{p} < .01$. Only 25 percent of subjects given the standard rule selected NOT-Q, in contrast to 62 percent of subjects given the explicit rule. This effect appears to be driven primarily by the higher rate of correct performance in the explicit rule conditions.

Qualitative Analysis of Reasons

This analysis will summarize the frequency and major types of reasons subjects gave for the standard and explicit rules. Most of the reasons subjects gave for their selections in the standard rule condition can be grouped into three general categories that roughly correspond to the three most common response categories (P alone, P & Q, and P & NOT-Q). In addition, some of the P alone and P & Q selections showed indications of a matching bias. Samples of the different types of reasons can be found in Appendix B.

Two of the 12 subjects in the standard rule condition gave reasons that indicated an interpretation that is consistent with material implication. Both subjects selected the logically correct P & NOT-Q. They wrote that the vowel had to have an even number on the other side and that the odd number could not have a vowel on the other side.

The three subjects selecting P & Q gave reasons indicating that for the rule to be true the A must have an even number on the back and the 8 must have a vowel on the back. An additional four subjects selected P alone and gave reasons stating that the A card must have an even number for the rule to be true.

Statements that dismissed cards because they were not mentioned in the rule were considered as evidence of matching. For example, one of the matching subjects wrote that "the rule said nothing at all about odd numbers, so we don't care what is on the other side" as a reason for not selecting the NOT-Q. There was evidence of matching in the reasons given by three of the subjects described earlier. Two selected P alone and one selected P & Q. In addition to these three matching subjects, one subject, who selected P & Q, did not even give reasons for the rejection of NOT-P and NOT-Q.

Evidence of task misinterpretations was found in the reasons given by three of the subjects. One subject selected P & NOT-P and gave reasons that indicated a mistaken belief that if the rule was true for one vowel/even number card and one consonant/odd number card, then it would be true for all of them. The reasons given by two subjects implied that they had interpreted the rule as a given rather than as a rule whose truth needed to be tested. Reasons for excluding a card from selection such as "obviously there is a vowel on the opposite side" (this reason was given for <u>not</u>

selecting Q) were taken as evidence that the rule was being interpreted as a given. One subject selected NOT-Q alone saying the rule might be wrong, but indicated that the rule was a given for the other three cards. The other subject interpreting the rule as a given did not select any of the cards.

For the explicit rule, eight of the 12 subjects giving reasons made the logically correct selection of P & NOT-Q. As with the standard rule, the reasons provided stated that the vowel would have to have an even number and the odd number would have to have a consonant for the rule to be true. Several subjects also wrote that if a vowel was found on the other side of the odd number, then this would falsify the rule.

None of the subjects selected P & Q. One subject gave a typical P-alone reason but selected P & NOT-P. The reason given for selecting NOT-P indicated that it really did not matter what was on the other side.

Three subjects gave reasons that either showed evidence of misunderstanding the task or were not understandable. For the most part, these selections and reasons were idiosyncratic. One subject treated the rule as a given, selecting Q alone. The other two subjects gave reasons that were uninterpretable because of their brevity and incompleteness.

In summary, the standard rule showed evidence of eliciting a matching strategy for several subjects. This

was not the case for the explicit rule. There was no evidence of matching in the reasons given in the explicit rule conditions. Overall, those subjects who made logically correct selections gave evidence of understanding the logical relationship being tested. For both types of rules, there was a substantial minority of subjects who misinterpreted the task to the extent that they could not have reached the logically correct solution. The most common misinterpretation was for the subject to interpret the rule as a given and not as a rule needing to be tested.

Table 3-1
Percents Correct and Logic Index Scores
for Experiment 2

	Reasons			
	Absent		Present	
Rule Type	Percent Correct	Logic Index	Percent Correct	Logic Index
Standard	0	0.00	17	0.75
Explicit	33	1.08	67	1.08

Note. *N* = 12 for each condition.

Table 3-2
Matching Index Scores for Experiment 2

	Reasons	
Rule Type	Absent	Present
Standard	1.33	0.75
Explicit	0.75	-0.08

Note. *N* = 12 for each condition.

Table 3-3
Frequency of Selection Combinations for Experiment 2

	Rule Type			
	Standard		Explicit	
Selection	Reasons Absent	Reasons Present	Reasons Absent	Reasons Present
P	1	4	3	0
NOT-P	0	0	0	0
Q	1	0	0	1
NOT-Q	0	1	1	0
P, NOT-P	0	1	0	1
P, Q	7	3	3	0
P, NOT-Q*	0	2	4	8
NOT-P, Q	0	0	0	1
NOT-P, NOT-Q	0	0	0	1
Q, NOT-Q	0	0	0	0
P, NOT-P, Q	0	0	0	0
P, NOT-P, NOT-Q	0	0	0	0
P, Q, NOT-Q	0	0	1	0
NOT-P, Q, NOT-Q	0	0	0	0
ALL	3	0	0	0
NONE	0	1	0	0
TOTAL	12	12	12	12

*Correct Answer

Table 3-4
Proportions of Individual Card Choices for Experiment 2

	Rule Type			
	Standard		Explicit	
Selection	Reasons Absent	Reasons Present	Reasons Absent	Reasons Present
P	.92	.83	.92	.75
NOT-P	.25	.08	.00	.25
Q	.92	.25	.33	.17
NOT-Q	.25	.25	.50	.75

_N_ = 12 for each condition.

CHAPTER 4
EXPERIMENT 3

Griggs (1989) found facilitation of performance (80 percent correct responses) for a selection task that was modified to eliminate Margolis's (1987) hypothesized semantic and scenario ambiguities. However, the modifications also made matching the consequent impossible. In a second experiment, Griggs found that correct responding dropped to 25 percent when the rule was modified to allow matching of the consequent.

The rule explications used in Experiments 1 and 2 made mention of all four card categories. This made a matching strategy ineffective for differentiating between the cards. Experiment 3 looked at an explicit rule that would still allow the subject to use a matching strategy. The rule ambiguity/lack of analytic processing hypothesis predicts that correct performance would be reduced for an explicit rule that allows matching. P alone responses should be common since the conditional nature of the rule is explicit. Matching should result in an initial selection of P & Q. However, because the rule makes the conditional relationship explicit, some logical processing might also be involved. This would make the recognition of the irrelevance of Q more likely and would lead to a selection of P alone.

The reasons response format from Experiment 2 was included as a factor. Having to give reasons should reduce matching in the matching rule condition, because attention would be forced on each card and reasons would have to be provided. Also, correct performance should be higher for both matching and nonmatching rules when reasons were required.

Methods

Subjects

Sixty-four introductory psychology students at the University of Florida served as subjects. None had participated in the earlier experiments.

Design

Subjects were run in one large group with each subject being given a single selection task problem. Sixteen subjects were randomly assigned to each of the four conditions created by a 2 x 2 factorial design. The independent variables were Rule Format (matching or nonmatching), and Reasons (present or absent). Four card orders were used to control for card position effects. They were P, NOT-P, Q, NOT-Q; NOT-Q, Q, NOT-P, P; NOT-P, NOT-Q, P, Q; and Q, P, NOT-Q, NOT-P.

Materials

The reasons response format was the same as in Experiment 2. Both rules used were explicit, but the matching rule read:

> A card with a <u>vowel</u> on its letter side can only have an <u>even</u> <u>number</u> on its number side, but a card with an <u>even</u> <u>number</u> on its number side does not have to have a <u>vowel</u> on its letter side.

For the nonmatching rule, the explicit rule used in Experiment 2 was changed to more closely resemble the matching rule described above. The nonmatching rule read:

> A card with a vowel on its letter side can only have an even number on its number side, but a card with a consonant on its letter side can have either an even or an odd number on its number side.

Results and Discussion

The results of Experiments 1 and 2 revealed two factors that improve performance on the abstract selections task. The use of an explicit rule paired with requiring reasons led two-thirds of the subjects to make the correct selection. Experiment 3 examined whether subjects will revert to a matching strategy if one is made more salient by mentioning only two of the four values in the rule. Reasons were again required allowing further qualitative analysis.

Quantitative Analysis

The percents correct and logic index scores for Experiment 3 can be found in Tables 4-1. Two-way analyses of variance were performed on both measures. The analysis for correct responses revealed a main effect of Reasons, $F(1,60) = 4.06$, $p < .05$. In the reasons-present conditions, 37.5 percent of subjects made the logically correct

selection. When reasons were absent, only 15.6 percent made the logically correct selection.

Matching index scores can be found in Table 4-2. A two-way analysis of variance for matching index scores revealed no significant main effects or interactions.

The actual selection frequencies can be seen in Table 4-3. It was hypothesized that the selection of P alone would be higher for the matching conditions. When the nonmatching rule was used, 18.7 percent of subjects selected P alone. When the matching rule was used, P alone selections rose to 37.5 percent. However, this difference was only marginally significant [$X^2(1, N = 64) = 2.80$, $p < .10$].

The frequency of individual card choices can be found in Table 4-4. A two-way analysis of variance was performed for each of the four cards. The analysis of the P-choice data revealed a significant interaction between Rule Format and Reasons, $F(1,60) = 4.35$, $p < .05$. As can be seen in Table 4-4, when the nonmatching rule is used without reasons, 69 percent of subjects include P in their selections. When reasons were given, the percentage of subjects including P in their selections jumps to 94 percent. With the matching rule and reasons absent, 94 percent of the subjects include P in their selections. This percentage goes down slightly, to 81 percent, when reasons are present. There were no other significant main effects or interactions.

In general, the results of the quantitative analysis for this experiment are quite noisy. Contrary to expectations, the performance was very poor for the nonmatching rule when reasons were not given. As predicted, P alone selections were more common with the matching rule. However, this effect was only marginally significant. A number of unusual selections occurred including four subjects who selected NOT-P & Q. The reasons given by the two subjects in the reasons-present conditions indicated the NOT-P & Q response may reflect a logical response to a misunderstood task. The next section will look in more detail at the reasons subjects gave and what these reasons may reveal about the subjects' interpretation of the task.

Qualitative Analysis of Reasons

Out of the 16 subjects giving reasons with the nonmatching rule, 7 made the logically correct selection of P & NOT-Q. The reasons given by these subjects reflected a material implication interpretation of the rule. In addition, two other subjects gave reasons that reflected a logically correct analysis of the rule but did not make the correct selections. One selected NOT-P & Q and indicated that the rule had been interpreted as a given not as a rule whose truth needed to be tested. This subject apparently misunderstood the nature of the task. The reasons given imply that this subject thought the task was to figure out what was on the other side of the cards. Since the rule defined what could be on the other side of the P and the

NOT-Q cards, they did not need to be turned over. The other subject picked all the cards even though the reasons claimed that either value could be on the back of the consonant and the even number. Two more subjects selected P, Q, & NOT-Q saying that the even number should have a vowel on the other side.

Two subjects selected P & Q and gave the typical reasons for these selections. Two other subjects selected P alone. One of these subjects recognized the irrelevance of the letter on the other side of the even number, but the other one failed to provide reasons for any of the cards that were not being turned over. Finally, one subject selected all the cards but gave reasons that were uninterpretable.

Of the 16 subjects giving reasons with the matching rule, 5 selected the logically correct P & NOT-Q. All but one of these gave reasons that reflected a clear logical understanding of the rule. The one exception failed to give a reason for the selection of the NOT-Q card. An additional subject gave reasons consistent with the P & NOT-Q choice but selected NOT-P & Q instead. The reasons given implied that the rule had been interpreted as a given not as a rule whose truth was in question.

Five subjects selected P alone. Four of the five gave reasons that indicated that the Q card could have any letter on the back. The remaining subject wrote that the Q card was "not part of the rule." An additional subject selected

NOT-P, Q, & NOT-Q but gave reasons consistent with the selection of P alone. The reasons clearly showed that the rule had been interpreted as a given. The reason for not selecting P is that the number would be even. The reason for selecting the other cards was that they could have either value on the other side. Another subject selected P, NOT-P, Q but indicated Q did not have to have a vowel. The reasons for selecting the NOT-P card and the Q card implied that the subject had interpreted the second part of the rule to mean even numbers had to have <u>both</u> vowels and consonants.

There was some evidence of matching in the reasons that were given by the subjects described above. Two of the subjects selecting P alone gave reasons that showed the use of a matching strategy in the elimination of at least some of the cards. For example, one of these subject rejected NOT-Q "because odd numbers are not mentioned in the rules." Two other subjects, not previously described, showed evidence of using a matching strategy. One choose Q alone and interpreted the rule as a given, not as a rule whose truth was under question. The other picked P & NOT-P, but gave evidence of rejecting NOT-Q because of a matching strategy.

Four subjects gave reasons that were uninterpretable or indicated task misunderstandings. Three of these subjects treated the rule as a given, and their performance has been addressed above. The fourth subject, selecting P, NOT-P &

NOT-Q, gave reasons that were uninterpretable and incomplete.

In summary, there was evidence that some subjects reverted to a matching strategy when that was available. Four subjects dismissed at least one card because it was not mentioned in the rule. P-alone-type reasons were given by six subjects in the matching condition compared to only two in the nonmatching condition. There were no P & Q type reasons for the matching rule and only two for the nonmatching rule. In the matching rule conditions, 15 subjects selected P alone, while only 3 subjects selected P & Q. The low rate of Q selections may indicate that, although matching was being used to dismiss NOT-P and NOT-Q, the Q card underwent additional analysis leading to its dismissal as well. There is also evidence that subjects can make a logical interpretation of the rule and successfully reason about it, regardless of whether the rule allows matching. Overall, approximately 25 percent of the subjects made the logically correct selections. The reasons given by 9 of the 16 subjects in the nonmatching condition and 6 of the 16 subjects in the matching condition reflected a material implication interpretation of the rule. Finally, reasons indicated that some subjects misinterpret the nature of the task in both the matching and explicit rule conditions. These subjects tended to interpret the rule as a given rather than something in need of testing.

Table 4-1
Percents Correct and Logic Index Scores
for Experiment 3

	Reasons			
	Absent		Present	
Rule Format	Percent Correct	Logic Index	Percent Correct	Logic Index
Matching	25	0.94	31	0.69
Nonmatching	6	0.31	43	1.00

Note. N = 16 for each condition.

Table 4-2
Matching Index Scores for Experiment 3

Reasons	Rule Format: Nonmatching	Rule Format: Matching
Absent	0.56	0.81
Present	0.50	0.31

Note. N = 16, for each condition

Table 4-3
Frequency of Selection Combinations for Experiment 3

	Rule Format			
	Nonmatching		Matching	
Selection	Reasons Absent	Reasons Present	Reasons Absent	Reasons Present
P	4	2	7	5
NOT-P	0	0	0	0
Q	0	0	1	1
NOT-Q	2	0	0	0
P, NOT-P	1	0	0	1
P, Q	3	2	3	0
P, NOT-Q*	1	7	4	5
NOT-P, Q	2	1	0	1
NOT-P, NOT-Q	0	0	0	0
Q, NOT-Q	1	0	0	0
P, NOT-P, Q	0	0	0	1
P, NOT-P, NOT-Q	0	0	1	1
P, Q, NOT-Q	1	2	0	0
NOT-P, Q, NOT-Q	0	0	0	1
ALL	1	2	0	0
NONE	0	0	0	0
TOTAL	16	16	16	16

*Correct answer

Table 4-4
Proportion of Individual Card Choices for Experiment 3

	Rule Format			
	Nonmatching		Matching	
Selection	Reasons Absent	Reasons Present	Reasons Absent	Reasons Present
P	.69	.94	.94	.81
NOT-P	.25	.19	.06	.31
Q	.50	.44	.25	.25
NOT-Q	.37	.69	.31	.44

Note. N = 16 for each condition.

CHAPTER 5
EXPERIMENT 4

In Experiment 3, performance in both nonmatching conditions was lower than in the comparable conditions in Experiment 2. The rule wording in Experiment 3 was changed to more closely resemble the rule wording of the matching rule. Experiment 4 examined whether the difference in performance could be attributed to the different rule wordings. In addition, Experiment 4 introduced violation instructions in an attempt to improve performance.

As reviewed earlier, violation instructions have been reported to improve performance in thematic problems. Yachanin (1986) argues that true-false instructions make the selection task more difficult. With true-false instructions, the subject must keep track of two potential hypotheses: The rule is true and the rule is false. With violation instructions, the subject is given a rule that is true. There is no other competing hypothesis to consider. With the added cognitive load created by the true-false instructions, the subject is more likely to adopt a short-cut strategy such as matching or verification of the rule. However, this account does not explain why violation instructions have not substantially enhanced performance on the abstract task.

If, as has been argued earlier, the central difficulty in the abstract selection task is rule ambiguity, then the lack of enhanced performance begins to make sense. If the rule is ambiguous, then the violating instances may be ambiguous as well. Without a clear understanding of what constitutes a violating instance, the subject would be inclined to use a short-cut strategy. If much of the difficulty associated with the standard abstract task is indeed a reflection of rule ambiguity, then the lack of a violation instructional effect for the standard abstract task would be expected. Conversely, if the rule is made explicit, then we may see an instructional effect for the abstract task.

Violation instructions may produce the additional benefit of eliminating some of the task misinterpretation found in Experiments 2 and 3. In those experiments, some subjects gave reasons that indicated they had taken the rule as a given, not as a rule whose truth was in question. By focusing attention on violating instances, the violation instructions may also serve to emphasize that the rule might not be followed.

Methods

Subjects

Sixty-four students from introductory psychology classes at the University of Florida served as subjects to partially fulfill a course requirement. None of the subjects had participated in the previous experiments.

Design

A 2 x 2 factorial design was used. The two factors were Rule Format (Experiment 2 explicit rule versus Experiment 3 explicit rule) and Instructions (true-false versus violation). Subjects were run in small groups with each subject given one selection task problem. Sixteen subjects were randomly assigned to each of the four conditions. Four card orders were used to control for card position effects. They were P, NOT-P, Q, NOT-Q; NOT-Q, Q, NOT-P, P; NOT-P, NOT-Q, P, Q; and Q, P, NOT-Q, NOT-P.

Materials

The two rules used were the explicit rules from Experiment 2 and Experiment 3. The Experiment 2 rule read:

> A card with a vowel on it can only have an even number, but a card with a consonant on it can have either an even or an odd number.

The Experiment 3 rule read:

> A card with a vowel on its letter side can only have an even number on its number side, but a card with a consonant on its letter side can have either an even or an odd number on its number side.

The true-false instructions were as follows: "Your task is to decide which card or cards <u>must</u> be turned over in order to find out whether the rule is true or false." To create the violation version the phrase "whether the rule is true or false" was changed to "whether or not the rule is

being violated." The reasons response format was used in all conditions.

Results and Discussion

In the first two experiments, performance was enhanced by the use of an explicit rule and a reasons response format. In the third experiment, some evidence of matching was found when matching was a viable strategy. However, correct performance in the nonmatching conditions was lower than in the comparable conditions from Experiment 2. Experiment 4 attempted to resolve this descrepancy and introduced a third potential facilitating variable, violation instructions; which was expected to increase correct performance.

Quantitative Analysis

Percents correct and logic index scores for all four conditions can be found in Table 5-1. No matching index scores were calculated since both rules had a nonmatching format. Two-way analyses of variance for correct responses and logic index scores revealed no evidence of differing performance for the two rule formats.

There was a significant main effect of Instructions for correct responses, $F(1,60) = 12.37$, $p < .001$, and logic index scores, $F(1,60) = 6.51$, $p < .05$. With true-false instructions, 37.5 percent of the subjects made the logically correct selection. This replicates the poorer performance found in Experiment 3. However, with violation instructions, 78 percent of the subjects were logically

correct. This is among the highest levels of correct performance that have been reported in the abstract selection task.

The frequency of selection combinations can be seen in Table 5-2. In the true-false conditions, the error pattern is varied. No single selection dominates the pattern, and no more than three subjects selected any particular erroneous selection combination. In the violation conditions, most of the subjects selected P & NOT-Q. Of the seven subjects who did not pick P & NOT-Q, two selected the partial insight combination P, Q, & NOT-Q recognizing the importance of NOT-Q but missing the irrelevance of Q; and two gave reasons that clearly showed they had interpreted the rule structure correctly but had misunderstood the task (see the qualitative analysis for further elaboration).

The proportion of individual card choices can be seen in Table 5-3. Two-way analyses of variance for the individual card choices revealed no significant main effects of Rule Format and no significant interactions. However, the analyses did reveal several significant main effects of Instructions. There was a significant main effect of Instructions on the proportion of subjects selecting NOT-P, $F(1,60) = 5.61$, $p < .05$. With true-false instructions, 28 percent of the subjects selected NOT-P. With violation instructions, it fell to only six percent. There was also a significant main effect of Instructions on the proportion of subjects selecting Q, $F(1,60) = 4.49$, $p < .05$. When the

true-false instructions were given, 47 percent of the subjects chose Q. When the violation instructions were given, it fell to 22 percent. Finally, there was a significant main effect of Instructions on proportion of subjects selecting NOT-Q, $F(1,60) = 6.23$, $p < .05$. With the true-false instructions, 66 percent of the subjects selected NOT-Q. When the violation instructions were used, it rose to 91 percent.

Qualitative Analysis of Reasons

Because there was no effect of Rule Format in the quantitative analysis, the reasons given for the two rule statements were combined for the qualitative analysis. In the true-false instruction condition, 12 of 32 subjects made the logically correct selection. The reasons these subjects gave indicated that the rule had been interpreted as material implication. Three other subjects gave reasons consistent with a P & NOT-Q choice, but they selected NOT-P & Q instead. These three subjects gave reasons showing that they had treated the rule as a given not as a rule whose truth was in question. Two other subjects selected P, NOT-P, & NOT-Q. They gave logically correct reasons for selecting the P card and the NOT-Q card. One of them clearly believed that for the rule to be true, consonants would have to have both even and odd numbers. The other subject's reason for selecting NOT-P is unclear.

Three subjects selected P, Q, & NOT-Q and gave reasons that maintained the need for a vowel on the back of the Q-

card. However, they gave logically correct reasons for their selections of P and NOT-Q. An additional subject selected all four cards but gave reasons consistent with a P, Q, & NOT-Q selection.

Three subjects selected P & Q with their reasons being consistent with the typical P & Q response. A fourth subject gave P & Q type reasons but selected NOT-P & NOT-Q. This subject treated the rule as a given, not as a rule whose truth was in question. Two additional subjects selected P alone and gave reasons that were typical of that selection.

Several subjects gave reasons that were either uninterpretable or revealed a misinterpretation of the task. In addition to the four subjects already described who treated the rule as a given, there were two others whose reasons fell in that category. These subjects selected Q alone. Two additional subjects selected all the cards: One treated the rule as a biconditional, the other changed the nature of the task saying that if there is no limit on the number of cards, then they should all be turned over. Finally, one subject selecting P & Q gave reasons that were incomplete and uninterpretable.

In the violation conditions, all but seven subjects made the logically correct P & NOT-Q selection and gave appropriate reasons. Of those seven subjects, two selected P, Q, & NOT-Q and gave the typical reasons. Another two subjects treated the rule as a given and selected NOT-P & Q.

One subject selected P & Q giving the typical reasons for such a selection. Two subjects selected Q & NOT-Q. One of them gave reasons that showed some evidence that the rule was taken as a given, though not consistently. The reasons for the final subject were uninterpretable.

When provided with violation instructions, subjects gave reasons for selecting NOT-Q that were different from those most commonly given with true-false instructions. With true-false instructions the typical reason to turn over the NOT-Q would be to make sure there is a consonant on the other side. With violation instructions, a few subjects still gave this reason; but more of them indicated they were making sure that there was not a vowel on the other side. While this is simply two different ways of describing the same restriction, the emphasis may be revealing. The violation instructions tended to elicit a falsification/violation strategy checking for values that do not conform to the rule. The true-false instructions, on the other hand, tended to illicit a verification strategy checking for values that do conform to the rule. Surprisingly, there was no difference between the reasons given for the two instruction types for selecting P. Verification type reasons predominated for both types of instruction.

In summary, correct performance was much more common when the violation instructions were used. Performance with the true-false instructions was more varied with the correct

response as the most common selection but no clear pattern in the selections or the reasons given for the many errors. When subjects were presented with the true-false instructions, a substantial amount of task misinterpretation occurred. Interpreting the rule as a given rather than as a rule whose truth is in question seemed to be the most common of these. Three subjects made a similar misinterpretation when the violation instructions were used.

Replications

The excellent performance for an abstract selection task problem using violation instructions is a new finding. In the violation-instructions conditions, the correct response rate was 78 percent. In previous studies of instructional effects, facilitation for violation instruction has not been reported for abstract problems. Therefore, two replications of the violation instructions condition were conducted. The first was an exact replication of the Experiment-2-explicit-rule/violation-instructions condition of Experiment 4. Sixteen subjects were run in small groups with each subject receiving one problem. The correct response rate was 69 percent. Three of the five subjects, making erroneous selections, selected the partial insight combination, P, Q, & NOT-Q. One subject selected NOT-P & Q. This subjects' reasons indicated that the rule had been interpreted as a given rather than as a rule that could be violated. The remaining subject selected P & NOT-P, and the reasons given were idiosyncratic. Because

the dominant error in the replication data was selecting P, Q, & NOT-Q, a second replication was conducted with violation instructions that would force the subject to limit their selection to two cards.

Margolis (1987, p. 155) proposed that changing the instruction statement to "Figure out which two cards could violate the rule and circle them" would lead subjects to select NOT-P & NOT-Q in the abstract selection problem. Griggs and Jackson (1990) tested this claim and found that with the Margolis instructions, 65 percent of subjects made the usually rare NOT-P & NOT-Q selection. However, in a second experiment, they introduced the Margolis instruction into an abstract permission selection task. Cheng and Holyoak (1985; Experiment 2) had previously found facilitation of performance (55 percent correct selections) for the abstract permission selection task with standard violation instructions. Griggs and Jackson (1990) found that the Margolis instructions led to even better performance (79 percent correct selections) although the difference was not significant. In the present study, the second replication investigated the effect of the Margolis violation instructions with an explicit rule.

The improved performance found by Griggs and Jackson (1990) led to the hypothesis that performance would also improve in the present explicit rule problem. Facilitation might occur in two different ways. Because only two cards were to be selected, subjects could no longer select P, Q, &

NOT-Q. Since the instructions lead to an increase in NOT-Q selections, it is likely that Q would be the card excluded. The second most common error was the interpretation of the rule as a given rather than a rule that could be violated. Since the Margolis instructions emphasize the cards that could violate the rule, taking it as a given should be reduced. The instructions were changed from the violation instructions used in Experiment 4 to read:

> Your task is to figure out which two cards could violate the rule. Indicate these cards by placing checkmarks in the appropriate places.

The phrase "Turn Over?" next to each card was replaced with the phrase "Could Violate?" Twenty-four subjects were run in small groups with each subject receiving one selection task problem.

The results for one subject were thrown out because the subject did not follow instructions by failing to select any cards. Of the remaining 23 subjects, 87 percent made the logically correct selection. However, the change in the instructions did not reduce the incidence of subjects interpreting the rule as a given. The remaining 13 percent of the subjects selected NOT-P & Q, and their reasons indicated that they had interpreted the rule as a given rather than a rule that could be violated.

In summary, the results of Experiment 4 and the replications reveal a reliable high rate of correct performance for the explicit-rule/reasons-present selection

task when violation instructions are used. The highest rate occurred when the Margolis violation instructions, which restricted selections to two cards, were used. None of the instructional manipulations, however, seemed to eliminate the infrequent but persistent interpretation of the rule as a given rather than as a rule that could be violated.

Table 5-1
Percents Correct and Logic Index Scores
for Experiment 4

	Rule Format			
	Exp. 2		Exp.3	
	Explicit Rule		Explicit Rule	
Instructions	Percent Correct	Logic Index	Percent Correct	Logic Index
True-False	44	0.75	31	0.69
Violation	81	1.37	75	1.62

Note. N = 16 for each condition.

Table 5-2
Frequency of Selection Combinations for Experiment 4

	Instructions			
	True-False		Violation	
Selection	Exp 2 Rule	Exp 3 Rule	Exp 2 Rule	Exp 3 Rule
P	1	1	0	0
NOT-P	0	0	0	0
Q	1	1	0	0
NOT-Q	0	0	0	0
P, NOT-P	0	0	0	0
P, Q	1	3	1	0
P, NOT-Q*	7	5	13	12
NOT-P, Q	2	1	2	0
NOT-P, NOT-Q	1	0	0	0
Q, NOT-Q	0	0	0	2
P, NOT-P, Q	0	0	0	0
P, NOT-P, NOT-Q	0	2	0	0
P, Q, NOT-Q	2	1	0	2
NOT-P, Q, NOT-Q	0	0	0	0
ALL	1	2	0	0
NONE	0	0	0	0
TOTAL	16	16	16	16

*Correct answer

Table 5-3
Frequency of Individual Card Choices for Experiment 4

	Instructions			
	True-False		Violation	
Selection	Exp 2 Rule	Exp 3 Rule	Exp 2 Rule	Exp 3 Rule
P	12	14	14	14
NOT-P	4	5	2	0
Q	7	8	3	4
NOT-Q	11	10	13	16

CHAPTER 6
GENERAL DISCUSSION

The results of Experiment 1 provide support for the contention that rule ambiguity contributes to the difficulty of Wason's selection task. Performance with an explicit rule was significantly better than performance with the standard rule. However, even with the explicit rule, performance continued to be poor (31 percent correct). The majority of subjects continued to make incorrect selections.

The results of Experiment 2 indicated that in addition to an explicit rule, providing reasons for the selection or nonselection of each card affected performance. The percent of subjects making the correct selections was higher for the explicit rule conditions than for the standard rule conditions (50 percent versus 8 percent). Providing reasons also produced a higher rate of correct responses (42 percent for the reasons-present conditions versus 17 percent for the reasons-absent conditions). In the explicit-rule/reasons-present condition, performance was at its highest, 67 percent correct.

In the standard-rule conditions there was evidence of matching. The matching response, P & Q, was the most common error with 42 percent of the subjects making this response. In the standard-rule/reasons-present condition, 3 of the 12

subjects gave reasons indicating the use of a matching strategy to dismiss from consideration those cards not mentioned in the rule. None of the subjects in the explicit-rule/reasons-present condition gave reasons indicating a matching response and the rate of P & Q selection across both explicit rule conditions was only 12 percent.

In both conditions requiring reasons, evidence was obtained indicating that some subjects had misinterpreted the rule. Six subjects gave reasons that either indicated a misinterpretation or were uninterpretable due to brevity or incompleteness. Although some of these responses were idiosyncratic, three subjects indicated that they had interpreted the rule as a given truth rather than a rule whose truth was to be determined.

In Experiment 3, the possibility that, if given the opportunity, subjects might revert to a matching strategy was considered. The results of this experiment were varied. The reasons provided by some of the subjects given the matching explicit rule indicated matching. As was predicted, P alone was the most common response for subjects given the matching rule. This is consistent with the use of a matching strategy to eliminate NOT-P and NOT-Q and the application of the explicit knowledge provided about rule structure to eliminate Q from the selection. Performance was poor in the nonmatching-rule/reasons-absent condition. Performance improved in the reasons-present condition, but

did not attain the level found in Experiment 2. Again, the reasons given provided evidence that a number of subjects had misunderstood the task. As in Experiment 2, four subjects took the rule as a given truth rather than a rule whose truth is to be tested.

In Experiment 4, performance for the explicit, nonmatching rules used in Experiment 2 and 3 were compared. In Experiment 3, the wording had been altered to more closely resemble the wording of the matching rule. No significant difference in performance for the two rules was observed. However, a significant instructional effect was found. With violation instructions, very few subjects made selections other than P & NOT-Q (78 percent correct). Performance levels with true-false instructions resembled performance in Experiment 3 (37.5 percent correct). As in Experiments 2 and 3, misunderstandings of the task were evident in the reasons given by a number of the subjects, most commonly accepting the rule as a given. The unusually high level of correct performance with violation instructions was replicated twice. The first was an exact replication of the Experiment-2-explicit-rule/violation-instructions condition in Experiment 4. Performance was slightly lower but still maintained a high level of correct responses (69 percent). The second replication used Margolis's instructions which require the subject to select two cards. This eliminated the P, Q, & NOT-Q selections

observed in Experiment 4 and drove the correct response rate to 87 percent!

The results of this series of experiments are consistent with the rule ambiguity/lack of analytic processing hypothesis. Providing an explicit rule structure was a necessary but not a sufficient condition for most subjects to produce logically correct performance. The introduction of factors that would increase analytic processing, especially of the NOT-Q card led to further increases in correct performance.

When reasons were not required, the explicit rule improved performance; but the majority of subjects still did not make the logically correct selection. P alone and P & Q selections dominated the error pattern. However, a substantial minority made atypical selections. Many of these atypical selections can be traced to misunderstandings about the nature of the task. Some of the misunderstandings were the result of interpreting the rule as a given and not as a rule whose truth was in question, while others were idiosyncratic.

The normally rare NOT-P & Q selection appeared at low levels in every experiment in this series. Although this appears to be a bizarre error and is treated as the worst possible response by the logic index, the reasons subjects gave for this response indicated that they had interpreted the rule as a given. In this light, the response reflects an accurate logical understanding of the rule structure.

Subjects who interpret the rule as a given see their task as identifying the cards that have an undetermined value on the back. If the rule were a given, those cards would be the NOT-P and the Q cards. Taking the rule as a given establishes that P must have Q on the other side, and NOT-Q must have NOT-P on the other side. Because the NOT-P and Q cards can have any value on the other side, the backs of these cards are unconstrained. Therefore, given this interpretation, NOT-P and Q are the cards that must be selected. This points out the importance of determining how the subject perceives the task when trying to understand the subject's card selections. This may also help explain the findings of an earlier study.

Yachanin and Tweney (1982) reported that 17 percent of subjects given an abstract selection task with an If P, Q rule selected NOT-P & NOT-Q. The only response combination that occurred more frequently was P & Q (50 percent). Yachanin and Tweney attribute the NOT-P & NOT Q selections to a "contorted disconfirmation strategy" (p. 89). However, the results of the present study suggest that Yachanin and Tweney's subjects may have misinterpreted the task. Most of Yachanin and Tweney's subjects interpreted the rule as constraining the contents of the P and Q cards, evidenced by the selection of P & Q. If some subjects understood the task to be determining what values actually appeared on the other side of the cards, as the subjects who took the rule

as a given did in the present experiment, then they would pick the two unconstrained cards, NOT-P and NOT-Q.

In the present study's Experiments 1 and 2, the standard rule produced its typical response pattern of mostly P & Q or P alone selections. Evans (1984) has argued that these responses are the result of a preattentive matching bias that results in only the P card and the Q card being judged relevant because they are the only ones mentioned in the rule. However, the present study provides some evidence that the matching bias, at least in some cases, may reflect an active strategy rather than a preattentive bias. In Experiment 2, several subjects making these typical selections gave reasons that indicated the use of a matching strategy. If it were a preattentive process, it is unlikely that it would be available for the subject to present as a reason. The results of Experiment 3 provided additional evidence of matching when an explicit rule mentioning only P and Q was examined.

In all of the experiments, performance for explicit rules was superior to typical performance on an abstract selection task. There are two ways that explicit rules may facilitate selection task performance. First, the explication eliminates ambiguity in the meaning of the rule. In Experiment 1 when the explication was added to the standard rule, performance improved. Correct performance was at its highest, however, when the standard rule was replaced with an explicit rule. The second benefit from the

use of an explicit rule is that it makes matching a useless strategy for differentiating between the cards. The presence of the standard rule allowed the subjects to continue to use a matching strategy leading to poorer performance. When the explicit rule was alone, matching was no longer possible and performance improved.

Informal observation of subjects' behavior revealed a common pattern of response in the reasons-present conditions. Initially subjects would mark all of their selections and then go back to write the reasons for them. While writing the reasons, subjects would stop, look back at the rule, and then sometimes erase the selections they had made and change them to the logically correct ones. While no record was kept of how many subjects followed this pattern, it was common. Therefore, subjects' reasons may sometimes not reflect their initial selection processes. However, because it is in the reason-writing stage that answers are changed, it seems safe to conclude that it does reflect the analytic processes underlying the final set of selections.

Johnson-Laird and Byrne's (1991) mental models theory may be a useful framework in which to understand how the reason-giving procedure impacts the selection process. They argue that subjects reason by constructing mental models representing possible states of affairs. The formation of these models is guided by the context, content, and type of reasoning task. However, these models are not necessarily

complete. Subjects only consider instances that are represented in their models. In the selection task, the models represent pairings of card values. At the model evaluation stage, subjects select the cards that are represented in their model and bear on the truth or falsity of the rule. Johnson-Laird and Byrne (p. 81) claim that reason-giving should make subjects more likely to envisage all possible states of affairs. The observation that subjects frequently change their selections while writing their reasons fits within this framework.

According to the mental models approach, subjects working on the selection task form an incomplete model where P is exhaustively paired with Q. That means that there is no possibility of P paired with any value other than Q in the model. If Q is also exhaustively paired with P, then the subject selects P & Q. If Q is not exhaustively paired with P (i.e., the possibility of Q paired with a value other than P is represented in the model), the subject selects P alone. In writing the reason for turning the P card over, there is an opportunity for NOT-Q to be brought to their attention. Subjects then reassess their model and realize the need to flesh it out so that it includes the NOT-Q card. The model may also get fleshed out as they consider reasons for not including NOT-Q in their initial selection. With the model now fleshed out to include the violating instance, P & NOT-Q, subjects can make the correct selection since

they recognize that this is the instance that bears on the truth or falsity of the rule.

The mental models approach is also useful in understanding the effect of violation instructions found in Experiment 4. Violation instructions make it more likely that the violating instance, P & NOT-Q, will be explicitly represented in the model. Several studies, described in the introduction, have found evidence that subjects will look for cards that would verify the rule--verification bias. In mental models terms, if subjects are more likely to look for verifying instances, then they are more likely to represent verifying instances in their model. Therefore, anything that might help overcome verification bias might be expected to improve performance by increasing the likelihood that nonverifying instances would be included in the model. If the task is to find instances that might violate the rule, then a subject should be drawn to represent in the model what a violation would entail. However, with true-false instructions, subjects can more easily opt to represent only verifying instances of the rule. Therefore, violation instructions should be more likely to lead to a fleshed out model.

The reasons given in the present experiments provided support for this explanation. Specifically, throughout these experiments the most common type of reason for turning over the vowel was to make sure that there was an even number on the other side. The equally accurate reason, to

make sure there was not an odd number, was uncommon. Likewise, the most common reason for turning over the odd number was framed in terms of verifying that there was a consonant, rather than checking for the presence of a vowel. However, when subjects were given violation instructions in Experiment 4, more than half the subjects made some mention of the fact that the presence of a vowel on the other side of the odd number would violate the rule. The violation instructions may improve performance by helping to overcome the verification bias. Surprisingly, subjects continued to give verification responses for the vowel card.

The effect of explication can also be integrated into the mental models account. Because the rule is unambiguous, subjects are less likely to search for short-cut strategies in building their models. In addition, since all the cards are mentioned in the rule, the most common short-cut strategy, matching, is rendered ineffective. As a result, subjects are more likely to fully develop their model and include NOT-Q in it.

None of the other theoretical perspectives outlined earlier can account for all of the findings in the present study. Cheng and Holyoak's (1985) pragmatic reasoning schema theory was designed to explain thematic content effects. It might be argued that the violation instructions trigger some sort of violation detection schema. In addition, rule explication might make a schema more salient since it clarifies the interpretation. However, it is hard

to imagine how reason giving could lead to the adoption of a different schema. Thus, it is unlikely that these effects are schema-driven.

Margolis (1987) can account for rule explication effects because the explicit rule eliminates what he calls semantic ambiguity. He might also argue that violation instructions led subjects to a closed scenario interpretation decreasing scenario ambiguity. However, reason-giving should have no impact on either semantic ambiguity or scenario ambiguity. Reason-giving does not change the rule statement so semantic ambiguity remains unchanged. In addition, there is no basis for expecting reason-giving to help define the scenario as closed rather than open. Therefore, the reasons effect remains unexplained in Margolis's theory.

Evans's (1984; 1989) two-stage model is the most likely candidate as an alternative explanation. During the initial heuristic stage, a relevance judgment is made using preattentive processes as a guide. In the second stage, those cards that were judged as relevant can be subjected to analytical processing. The explicit rule effect could be linked to blocking of matching bias. However, Evans maintains that performance on the selection task is determined at the heuristic stage. He does not believe that analytical processing is involved in selection task performance. With this position, it would be difficult to account for the effect of reasons seen in the present study.

Therefore, the theory would have to be modified to allow second stage processing in order to account for all of the present results. In addition, since the analytical stage of his model has remained so underspecified, the mechanisms that might account for the reason-giving effect have not been defined. Thus, while it might be able to explain the present results with some modifications or increased specificity, in its present state it cannot.

In summary, the present study found evidence for three major factors that influence performance in the selection task. The use of an explicit rule improved performance indicating an interpretational element in the difficulty of the standard task. However, it was also found that some subjects will still demonstrate a matching bias for an explicit rule when given the opportunity. Two types of instructional factors also played a role. Asking subjects to give reasons for their selections and the use of violation instructions led to improved performance. Performance was also better when subjects were given violation rather than true-false instructions.

While the present series of experiments makes clear the facilitation produced by these three factors, several questions still need further investigation. First, the roles of a forced-choice procedure and providing reasons need to be assessed as separate factors. In the present study these factors were combined in order to get reasons for nonselections as well as selections. However, mental

models theory suggests that forced choice would make a complete model more likely by forcing consideration of each card. Some pilot work has suggested an effect of forced choice, but the effect has not been reliable. Second, task misinterpretations need to be examined and addressed. Throughout the present study, a small but persistent percentage of subjects interpreted the rule as a given rather than a rule whose truth was in question despite all the manipulations used to improve performance. The cause of these misinterpretations needs to be systematically investigated. Finally, the effect of matching for explicit rules needs to be further investigated. In Experiment 3, some evidence of matching was found, but an investigation of factors that might mediate this effect is needed. For example, does matching continue when violation instructions are used rather than true-false instructions?

These research questions relate to three different critical aspects of the selection task: Subjects' interpretation of the task, the selection instructions, and subjects' analytical strategies. A more thorough understanding of these will allow us to understand the reasoning processes behind performance on the selection task. Too often theoretical explanations of selection task performance have focused on one of these aspects while ignoring the others. The mental models approach is the most promising theoretical framework for integrating the effects

due to all three factors; and therefore, it seems to be the most appropriate approach to guide further investigations.

APPENDIX A
THE BASIC SELECTION TASK

Each of the boxes below represents a card lying on a table. Each one of the cards has a letter on one side and a number on the other side. Here is a rule:

If a card has a vowel on its letter side, then it has an even number on its number side.

As you can see two of the cards are letter side up, and two of the cards are number side up. Your task is to decide which card or cards must be turned over in order to find out whether the rule is true or false. Circle only the card or cards that must be turned over.

A	K	8	5

APPENDIX B
SAMPLES OF THE MOST COMMON TYPES OF REASONS

This appendix contains sample reasons that are characteristic of the different types described in the main text of the present study. The values shown under the heading "Selection" are the potential cards the subject had to choose from. Each card that was selected is so indicated by an "x" being placed next to it. Next to each card, under the heading "Reason" is found the reason the subject gave for the decision made about that card regardless of whether it was selected or not selected.

Sample reasons for P & NOT-Q selections

Selection	Reason
A x	It must be even to be true.
5 x	It must be a consonant to be true.
K	A consonant can be even or odd so either way it is true.
8	A vowel or consonant can have even so it is true either way.

Sample reasons for P & Q selections

Selection	Reason
A x	To see if it is an even number
5	It can have either a vowel or a consonant.
K	It can have either number.
8 x	To see if it has a vowel on the other side

Sample reasons for P alone selections

Selection	Reason
A x	By turning over the vowel "A", there should be an even number on the other side according to the rule. If the number is odd, the rule is false.
5	If this card is turned over, any letter of the alphabet could appear (vowel or consonant).
K	By turning over the K, any number (even or odd) is possible.
8	If this card is turned over, any letter could appear.

Sample reasons for matching responses (matching used to reject 5 & K)

Selection	Reason
A x	We have to see if in fact there is an even # on the other side.
5	The rule said nothing at all about odd #'s, so we don't care what's on the other side.
K	The rule said nothing at all about consonants, so we don't care what's on the other side.
8 x	We have to see if in fact there is a vowel on the other side.

Sample reasons for rule-as-given (rule-as-given reasons given for rejecting A and 5)

Selection	Reason
A	Since its a vowel we know the # is even.
5	Let it stay--because its odd it must be a consonant so we don't need to determine what it is we already know.
K x	We need to know if the # is even or odd because the letter is a consonant.
8 x	Turn over because its an even # the letter could be either vowel or consonant.

REFERENCES

Beattie, J., & Baron, J. (1988). Confirmation and matching biases in hypothesis testing. The Quarterly Journal of Experimental Psychology, 40A, 269-297.

Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. Quarterly Journal of Experimental Psychology, 35A, 39-49.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. Cognitive Psychology, 17, 391-416.

Cheng, P. W., & Holyoak, K. J. (1989). On the natural selection of reasoning theories. Cognition, 33, 285-313.

Chrostowski, J. J., & Griggs, R. A. (1985). The effects of problem content, instructions, and verbalization procedure on Wason's selection task. Current Psychological Research & Reviews, 4, 99-107.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? The Behavioral and Brain Sciences, 4, 317-370.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Cognition, 31, 187-286.

Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. Memory and Cognition, 10, 496-502.

Dominowski, R. L. (1989, September). Success and failure on the four-card problem. Paper presented at the Sixth Annual Conference of the Cognitive Psychology Section, British Psychological Society, Cambridge, England.

Dominowski, R. L. (1990). Problem solving and metacognition. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie, & G. Erdos (Eds.), Lines of thinking: Reflections on the psychology of thought: Vol. 2. Skills, emotions, creative processes, individual differences and teaching thinking. (pp. 313-328). Chichester, England: John Wiley & Sons.

Evans, J. St. B. T. (1972). Interpretation and matching bias in a reasoning task. Quarterly Journal of Experimental Psychology, 24, 193-199.

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. British Journal of Psychology, 75, 451-468.

Evans, J. St. B. T. (1989). Bias in human reasoning: Causes and consequences. East Sussex: Lawrence Erlbaum.

Evans, J. St. B. T. (1992). Relevance and reasoning in the selection task. Manuscript submitted for publication.

Evans, J. St. B. T., & Lynch, J. S. (1973). Matching bias in the selection task. British Journal of Psychology, 64, 391-397.

Evans, J. St. B. T., & Wason, P. C. (1976). Rationalization in a reasoning task. British Journal of Psychology, 67, 479-486.

Fillenbaum, S. (1975). If: Some uses. Psychological Research, 37, 245-260.

Geis, M. L., & Zwicky, A. M. (1971). On invited inferences. Linguistic Inquiry, 2, 560-566.

Griggs, R. A. (1983). The role of problem content in the selection task and THOG problem. In J. St.B. T. Evans (Ed.), Thinking and reasoning: Psychological approaches (pp. 16-43). London: Routledge and Kegan Paul.

Griggs, R. A. (1984). Memory cuing and instructional effects on Wason's selection task. Current Psychological Research & Reviews, 3, 3-10.

Griggs, R. A. (1989). To "see" or not to "see": That is the selection task. Quarterly Journal of Experimental Psychology, 41A, 517-529.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. British Journal of Psychology, 73, 407-420.

Griggs, R. A., & Cox, J. R. (1983). The effect of problem content and negation on Wason's selection task. Quarterly Journal of Experimental Psychology, 35A, 519-533.

Griggs, R. A., & Cox, J. R. (1992). Permission schemas and the selection task. Manuscript submitted for publication.

Hoch, S. J., & Tschirgi, J. E. (1983). Cue redundancy and extra logical inferences in a deductive reasoning task. Memory and Cognition, 11, 200-209.

Hoch, S. J., & Tschirgi, J. E. (1985). Logical knowledge and cue redundancy in deductive reasoning. Memory and Cognition, 13, 435-462.

Jackson, S. L., & Griggs, R. A. (1988). Education and the selection task. Bulletin of Psychonomic Society, 26, 327-330.

Jackson, S. L., & Griggs, R. A. (1990). The elusive pragmatic reasoning schemas effect. Quarterly Journal of Experimental Psychology, 42A, 353-373.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). Deduction. London: Lawrence Erlbaum Associates Ltd.

Johnson-Laird, P. N., & Tagart, J. (1969). How implication is understood. American Journal of Psychology, 82, 369-373.

Johnson-Laird, P. N., & Wason, P. C. (1970a). A theoretical analysis of insight into a reasoning task. Cognitive Psychology, 1, 134-148.

Johnson-Laird, P. N., & Wason, P. C. (1970b). Insight into a logical relation. Quarterly Journal of Experimental Psychology, 22, 49-61.

Klaczynski, P. A., Gelfand, H., & Reese, H. W. (1989). Transfer of conditional reasoning: Effects of explanations and initial problem types. Memory and Cognition, 17, 208-220.

Lawson, A. E. (1987). The four-card problem resolved? Formal operational reasoning and reasoning to a contradiction. Journal of Research in Science Teaching, 24, 611-627.

Legrenzi, P. (1971). Discovery as a means to understanding. Quarterly Journal of Experimental Psychology, 23, 417-422.

Lunzer, E. A., Harrison, C., & Davey, M. (1972). The four-card problem and the generality of formal reasoning. Quarterly Journal of Experimental Psychology, 24, 326-339.

Marcus, S. L., & Rips, L. J. (1979) Conditional reasoning. Journal of Verbal Learning and Verbal Behavior, 18, 199-223.

Margolis, H. (1987). Patterns, thinking and cognition: A theory of judgment. Chicago, IL: The University of Chicago Press.

Markovits, H. (1984). Awareness of the "possible" as a mediator of formal thinking in conditional reasoning problems. British Journal of Psychology, 75, 367-376.

Markovits, H. (1985). Incorrect conditional reasoning among adults: Competence or performance? British Journal of Psychology, 76, 241-247.

Markovits, H. (1988). Conditional reasoning, representation, and empirical evidence on a concrete task. Quarterly Journal of Experimental Psychology, 40A, 483-495.

Mosconi, G., & D'Urso, V. (1974, October). The selection task from the standpoint of the theory of double code. Paper presented at the Conference Deductive Reasoning, University of Trento, Italy.

Oaksford, M., & Stenning, K. (1992). Reasoning with conditionals containing negated constituents. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 835-854.

Pollard, P. (1981). The effect of thematic content on the "Wason selection task." Current Psychological Research, 1, 21-29.

Pollard, P., & Evans, J. St.B. T. (1987). Content and context effects in reasoning. American Journal of Psychology, 100, 41-60.

Rips, L. J., & Marcus, S. L. (1977). Suppositions and analysis of conditional sentences. In M. A. Just and P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, NJ: Erlbaum.

Roth, E. M. (1979). Facilitating insight in a reasoning task. British Journal of Psychology, 70, 265-271.

Rumain, B., Connell, J., & Braine, M. D. S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: If is not the biconditional. Developmental Psychology, 19, 471-481.

Smalley, N. S. (1974). Evaluating a rule against possible instances. British Journal of Psychology, 65, 293-304.

Taplin, J. E. (1971). Reasoning with conditional sentences. Journal of Verbal Learning and Verbal Behavior, 10, 219-225.

Taplin, J. E., & Staudenmayer, H. (1973). Interpretation of abstract conditional sentences in deductive reasoning. Journal of Verbal Learning and Verbal Behavior, 12, 530-542.

Valentine, E. R. (1985). The effects of instructions on performance in the Wason selection task. Current Psychological Research & Reviews, 4, 214-223.

Van Duyne, P. C. (1976). Necessity and Contingency in reasoning. Acta Psychologica, 40, 85-101.

Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), New Horizons in Psychology. Harmondsworth, England: Penguin Books.

Wason, P. C. (1968). Reasoning about a rule. Quarterly Journal of Experimental Psychology, 20, 273-281.

Wason, P. C. (1969). Regression in reasoning? British Journal of Psychology, 60, 471-480.

Wason, P. C., & Johnson-Laird, P. N. (1972). Psychology of reasoning: Structure and content. London: Batsford. Cambridge, MA: Harvard University Press.

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. Quarterly Journal of Experimental Psychology, 23, 63-71.

Yachanin, S. A. (1986). Facilitation in Wason's selection task: Content and instruction. Current Psychological Research & Reviews, 5, 20-29.

Yachanin, S. A., & Tweney, R. D. (1982). The effect of thematic content on cognitive strategies in the four-card selection task. Bulletin of the Psychonomic Society, 19, 87-90.

BIOGRAPHICAL SKETCH

Richard D. Platt was born on August 18, 1962, in Newton, Kansas. In 1985, he received his Bachelor of Arts degree from Bethel College in North Newton, Kansas. He received his Master of Science degree from the University of Florida in May 1989. He anticipates completing his Ph.D. in December 1992.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard A. Griggs, Chairperson
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

C. Michael Levy
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ira S. Fischler
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

M. Jeffrey Farrar
Associate Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James Algina
Professor of Foundations of Education

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December, 1992

Dean, Graduate School